ENDEAVOUR

A quarterly review, published in five languages,
designed to record the progress of the sciences
in the service of mankind

VOLUME XX October 1961 NUMBER 80

CONTENTS

061.22(100):5
Editorial: An International Council of Sciences 175

547.473.2:577.15
The Organic and Biological Chemistry of Mevalonic Acid — A. F. WAGNER, A.B., M.A., Ph.D., and K. FOLKERS, B.S., Ph.D. — 177

550.311:551.241
The Mohole — SIR EDWARD BULLARD, M.A., Sc.D., F.R.S. — 188

612.843.3:535.7
Light Quanta and Vision — M. H. PIRENNE, M.A., Ph.D., Dr.Sc. — 197

92:552.3
Sir James Hall, Bt. (1761–1832) — V. A. EYLES, D.Sc., F.R.S.E., F.G.S. — 210

581.524.441:553.97:551.345
Surface Patterns in Boreal Peatlands — H. SJÖRS, Fil. dr. — 217

539.124
Positronium — F. F. HEYMANN, B.Sc.(Eng.), Ph.D.(Eng.). — 225

049.3
Book Reviews 232

Short Notices of Books 235

Notes on Contributors 236

Some Books Received .. .. Cover iii

Editor: TREVOR I. WILLIAMS, M.A., B.Sc., D.Phil., F.R.I.C.
Assistant Editor: G. N. J. BECK, B.Sc., A.M.I.E.E.
Editorial Assistant: W. A. OSMAN, B.Sc., M.Sc., A.R.C.S.

Imperial Chemical Industries Limited, Millbank, London, S.W.1
Communications to the Editor should be addressed to North Block, Thames House, Millbank, London, S.W.1

An international council of sciences

RECEIVED
11 OCT 1961
LIBRARY
WIDNES LABORATORY
WIDNES

Traditionally, science knows no frontiers, and a free exchange of ideas between men of science in different lands has always been looked upon as of the first importance. In the past, this freedom has transcended the bitterest national antipathies, even war itself. Completely free exchange is no longer practicable, but the principle remains. Today there is virtually no major field of science in which research is not being pursued in several different countries, and—to the advantage of all—new knowledge and new ideas are widely and quickly exchanged through the scientific journals, through conferences, or by private communication between individuals. For many purposes, such arrangements meet the needs of the day, and there are strong arguments against trying to establish a more rigid organization: in science, as in every human endeavour, much is sacrificed when the human touch is lost. There are, nevertheless, many instances of international co-operation in which a clearly defined allocation of responsibilities is not merely desirable but essential.

The ninth General Assembly of the International Council of Scientific Unions (ICSU), held in London last month, provided a timely reminder of the great and growing importance of the existence of formal machinery for international collaboration in certain kinds of scientific projects. The International Geophysical Year, organized by a special committee of ICSU, was by far the largest scheme of international co-operation in science that has ever been carried out. Sixty-seven countries took part and more than two thousand observing stations, among them fifty-seven in the Antarctic, were established. Although a scientific project of such magnitude can only occasionally be undertaken, it is a forceful indication of what can be achieved by joint international effort. No less important, it is a welcome reminder that close and friendly scientific co-operation is possible even between countries whose political differences seem irreconcilable.

The International Geophysical Year was of further significance in relation to ICSU; it was a powerful demonstration of the Council's ability to initiate a really great enterprise. It should therefore inspire the confidence necessary to secure adequate support for a further expansion of its activities. Such an expansion seems very necessary, for since the last war there has been an immense expansion in almost every branch of science, accompanied by increasing co-operation between different scientific disciplines and between scientists of different countries. It is clear that this fruitful trend will continue; moreover, the very existence of an efficient and adequately endowed organizing body will in itself do much to encourage it. ICSU does not itself collect and disburse funds for research. Its function is to co-ordinate the work of others and fit it into an effective pattern; it undoubtedly catalyses much research that would not otherwise be done.

ICSU has grown to its present stature from modest beginnings, that can be traced back to the second half of the nineteenth century, when there existed a number of associations of rather highly specialized natures. Some were organized inter-governmentally, but many were fostered by Germany; in 1900 they were brought together as the International Association of Academies. German influence was very strong, and in fact the Association was dominated by the academies of Berlin, Göttingen, Heidelberg, Leipzig, Munich, and Vienna. The first world war almost completely put a stop to its activities; afterwards, the International Research Council was founded on the joint initiative of the *Académie des Sciences*, Paris; the Royal Society of London; and the National Academy of Sciences, Washington. This Council served as a co-ordinator for various scientific unions, including chemistry, biological sciences, and physics; but originally it was no more truly international than the old Association of Academies, for it was really limited to the Allied countries. In 1931 the Council was reconstituted as the International Council of Scientific Unions, with aims and constitution far more in accordance with its title. Its defined purpose was to co-ordinate the work of the national adhering organizations and of the various international unions. Each of the last was given complete autonomy, and the Council's function approximated to that of a federal parliament, regulating what was common to all, but not interfering in the work of individual unions.

Again war disrupted the work, but soon afterwards the Council was revived once more. Today, it has two main aims. Firstly, to co-ordinate and facilitate the activities of the unions; secondly, to serve as a co-ordinating centre for the national organizations adhering to it. At present, the fourteen unions fall into two classes—general and

specialized—distinguished by their voting powers and representation. The former includes astronomy, geodesy and geophysics, pure and applied chemistry, pure and applied physics, and the biological sciences. Among the specialized unions are those for radio-science, crystallography, physiological sciences, biochemistry, and the history and philosophy of science.

While ICSU was being reconstituted, a new international agency interested in science appeared, the United Nations Educational, Scientific, and Cultural Organisation (Unesco). There are fundamental differences between them; whereas Unesco is an inter-governmental organization whose adherents are its member states, ICSU is an association of national academies of science or national research councils on the one hand, and of the international scientific unions on the other. Nevertheless, the interests of the two organizations naturally overlap, and, to avoid wasteful duplication, their respective fields of interest were defined in an agreement that has been renewed annually since 1946. Unesco provides ICSU with an annual subvention and is represented at the meetings of its Executive Board and of the General Assembly. In return, ICSU gives Unesco expert advice when called upon to do so.

The present activities of ICSU are too extensive for more than a few of the most important features to be summarized here; a full account was, however, recently written for this journal by the late Sir Harold Spencer Jones, formerly Secretary General (ENDEAVOUR, **18**, 88, 1959). Apart from organizing big international projects of observation and experiment, ICSU is active in co-ordinating scientific abstracting, and maintains a system of what are called permanent services.

The Council's interest in abstracts began in 1949, when Unesco held a conference in Paris on this subject. In the same year, following its General Assembly in Copenhagen, ICSU appointed a joint commission on physics abstracting, which in 1950 took over complete responsibility for this from Unesco. In 1951, the ICSU Abstracting Board was established, and now works in collaboration with 'Physics Abstracts', *Bulletin Signalétique*, *Physikalische Berichte*, and the Institute of Scientific Information, Moscow; its guiding principle is that for any particular branch of science there should be only one abstracting journal in each language. Subsequently this interest was extended to the field of chemistry, and today the Abstracting Board co-ordinates the work of the principal chemical abstracting journals in the English, French, German, and Russian languages.

The Federation of Astronomical and Geophysical Services was established in 1956. It now numbers eleven services in all, and serves the unions of astronomy, of geodesy and geophysics, and of scientific radio. The services include an international time bureau; a gravimetric bureau; an earth tide service; and a central ursigram service for the transmission of coded ionospheric, solar, geomagnetic, and other geophysical data.

In addition, ICSU has four special committees, dealing respectively with geophysics, oceanic research, Antarctic research, and space research. At present the most important task of the committee for geophysics is to ensure the fullest possible exploitation of the results of the International Geophysical Year, but already it is making plans for further comparable projects. These include a world magnetic survey—timed for 1964–66, when solar activity will be at its next minimum—and an exploration of the mantle of the Earth. Among the projects of the special committee on oceanic research is one for a detailed survey of the Indian Ocean in 1962–63; in this, at least twenty survey vessels, provided by ten nations, will take part. The Indian Ocean has been chosen for three main reasons. Firstly, it is the least known ocean of the world. Secondly, the winds and currents can be studied through a complete monsoon cycle, during which the direction of the prevailing wind is reversed. Thirdly, the fishing areas of this ocean are of great economic importance to the countries bordering upon it. In space research, the interest of the special committee, established in 1958, is in fundamental investigation, rather than in the technological problem of the design and launching of satellites and space probes.

The effective discharge of these diverse tasks demands adequate machinery and sufficient funds to maintain it. Hitherto its total income has been no more than £115 000 [$320 000]. That so much has been achieved for so little is due to the fact that most of the work is voluntary, though naturally travel and other incidental expenses must be paid; the permanent central directorate, at the Hague, is very small. Nevertheless, the time has clearly come when the organizational tasks ahead are greater than those ICSU can undertake with its present staff and resources. At a time when international expenditure on research is measured in hundreds of millions of pounds annually, it is very much to be hoped that ICSU will find no difficulty in obtaining the relatively very modest sum required to maintain and extend its services.

The organic and biological chemistry of mevalonic acid

A. F. WAGNER and K. FOLKERS

Since its identification in 1956, mevalonic acid has been recognized as a key substance in the biosynthesis of a wide range of very important natural products, including steroids, carotenoids, and terpenes. Its discovery is therefore of major scientific importance. This article, by pioneers in the field, reviews the research that led to the recognition of mevalonic acid as a growth factor, its isolation, the identification of its structure and configuration, its synthesis, and the elucidation of its role in a number of biosynthetic processes. The final section of the article makes it clear that, in addition, mevalonic acid is of fundamental biosynthetic importance in a number of fields that have so far been only briefly investigated.

DISCOVERY AND ISOLATION

Many investigators have used materials such as distillers' dried solubles for studying factors that promote the growth of animals. A. F. Novak, S. M. Hauge, and C. W. Carrick [1] reported in 1947 that such material contained an unidentified factor necessary for the growth of chicks, and in the following year a concentrate prepared from this source was designated 'vitamin B_{13}' [2]. Reports that a 'vitamin B_{13}' concentrate would promote growth in both chicks and rats stimulated studies in other laboratories.

In 1956, H. R. Skeggs, L. D. Wright, E. L. Cresson, G. D. E. MacRae, C. H. Hoffman, D. E. Wolf, and K. Folkers [3] subjected a 'vitamin B_{13}' concentrate to countercurrent distribution, and tested various fractions so obtained for their ability to promote microbial growth. Certain fractions were found to be highly active in promoting the growth, in an acetate-free medium, of *Lactobacillus acidophilus* (strain ATCC 4963). This activity was shown to be distinct from that of lipoic acid, a growth-promoting factor obtained from liver, and of all other compounds; these studies resulted in the discovery of a new acetate-replacing factor that was designated mevalonic acid. In these studies, the growth-promoting activity of 'vitamin B_{13}' concentrates in the chick was not confirmed; W. H. Ott and his colleagues [4] considered the growth tests with the concentrates to be negative, or at best equivocal in chicks. Purification of the new acetate-replacing activity was therefore continued using assays based upon the growth of *L. acidophilus* ATCC 4963 in an acetate-free medium.

To allay suspicion that the activity might be related to a known compound, or a combination of known compounds, over 150 compounds of biological significance were tested; not one showed comparable activity in promoting the growth of *L. acidophilus* in the absence of acetate. A demonstration that the factor was not a combined form of lipoic acid was particularly significant, because this compound had been shown to be present in concentrates. When microbiological assays for both lipoic acid and the 'new factor' were shown to be unchanged after hydrolysis with normal sulphuric acid, it was concluded that the material being studied contained no conjugate of lipoic acid that was acid-susceptible and unavailable to micro-organisms. The nearly quantitative recoveries of activity in various fractionation studies were also significant, since unaccountable losses in activity would have resulted if the microbial response had been due to a combination of compounds rather than to a single one. Further experiments with concentrates of the new factor demonstrated its ability to promote the growth of *Lactobacillus bifidus*, *L. bulgaricus*, and *Thermobacterium acidophilus*.

G. Tamura [5] detected a growth factor in Japanese rice-wine (sake) and also in culture media in which organisms such as *Aspergillus*, *Penicillium*, *Monilia*, *Willia*, and *Lactobacillus* had grown. The factor was designated hiochic acid because of its activity in promoting growth of the 'true hiochi bacteria', *Lactobacillus homohiochi* and *L. heterohiochi*. On the basis of the identical infrared spectra and microbiological activities of pure compounds obtained from both sources, G. Tamura and K. Folkers [6] later reported hiochic and mevalonic acids to be identical.

L. D. Wright and his colleagues [7] obtained highly purified concentrates of mevalonic acid from distillers' dried solubles. The active substance was extracted from the material with methanol, and this methanol-soluble fraction was partially purified by continuous extraction with chloroform. The chloroform-soluble portion was purified by chromatography (on Superfiltrol), and inert material in a series of selected eluates was precipitated with ethanol. When an aqueous solution of the concentrate was passed through a cation-exchange resin (Dowex-50), the active compound ran through rapidly, close behind the solvent front, and left behind many impurities. Selected fractions from this stage were then neutralized, and the factor was adsorbed on an anion-exchange resin (Dowex-1). Selective elution from the resin with 0·05N formic acid resulted in almost quantitative recovery of the activity. These two stages of ion-exchange purification achieved a five-fold concentration. After several successive countercurrent distributions of the richest fractions between the phases of an ethanol/water/chloroform mixture, the factor was obtained as a clear oil that weight-distribution and activity curves showed to be almost pure. About 150 mg of oil, obtained from 800 kg of starting material, were thus made available for the first structural studies.

STRUCTURE DETERMINATION

In aqueous solution, the newly isolated compound behaved as an acid; on potentiometric titration of an aqueous solution with alkali, the titration curve rose sharply and then drifted, in a manner characteristic of lactones. The infrared absorption spectrum of the compound also was compatible with a lactone function. In chloroform solution, absorption bands were observed that indicated the presence of hydroxyl and ester groups. In morpholine solution, the absorption band (5·78 μ) corresponding to the ester function slowly decreased in intensity during 48 hours, and a new band (6·1 μ), corresponding to a carboxylate ion, appeared and slowly increased in intensity. Such behaviour is typical of lactones, and, on the basis of the rate of ring opening, the function was assumed to be a δ- rather than a γ-lactone.

The highly purified oil was converted to the crystalline benzhydrylamide before further structural studies were undertaken. This derivative contained one C-methyl group and one hydroxyl group that could be acylated. Elemental analysis of the benzhydrylamide and its corresponding monoacetate indicated the molecular formula $C_6H_{12}O_4$ for the active compound, and $C_6H_{10}O_3$ for the corresponding lactone. Since spectral and titration data had already indicated the factor itself to be a δ-hydroxy acid, it appeared that placing a methyl group and one more hydroxyl group on the carbon chain would satisfy the elemental composition and complete the elucidation of the structure.

Neither the lactone nor its corresponding benzhydrylamide reacted with alkaline periodate or with hypoiodite, showing that the molecule does not contain an adjacent γ-hydroxyl group or

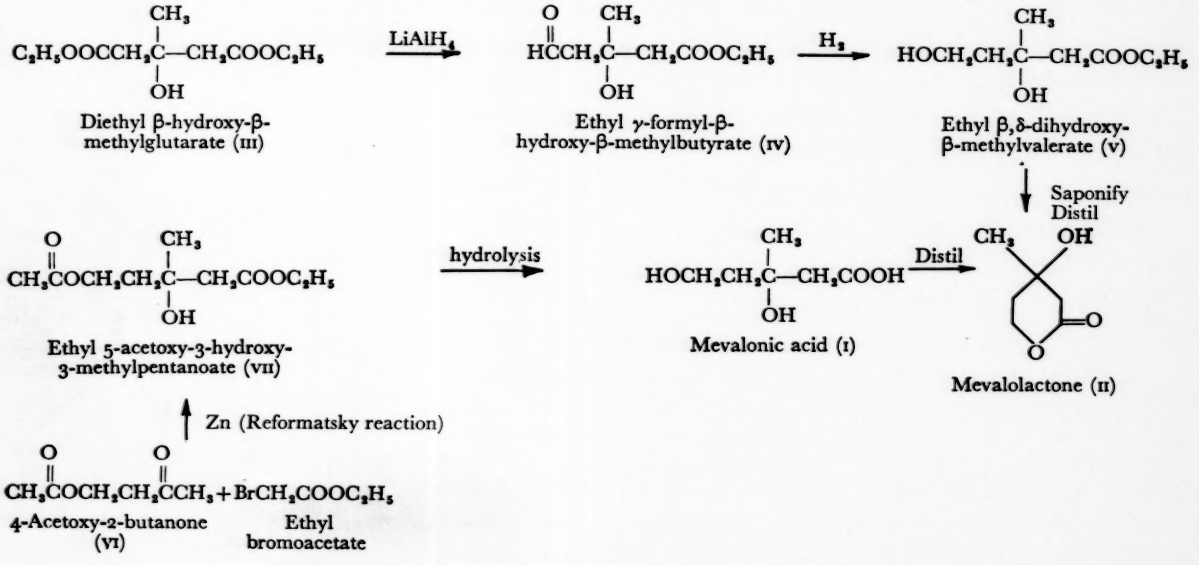


FIGURE 1 – *Original syntheses of mevalonic acid and mevalolactone.*

an α-hydroxyethyl substituent. Assignment of the hydroxyl group to the α-position was eliminated because of the pK of the acid; consequently, the β-position was the sole remaining possibility for the hydroxyl group.

The methyl group also was assigned to the β-position. Of the possibilities, the δ-position was eliminated as the site of the methyl group because the iodoform test was negative. Of the three remaining positions, the β-position was chosen because the behaviour of the compound on acetylation was best explained by the presence of a tertiary hydroxyl group.

On this basis, D. E. Wolf, C. H. Hoffman, P. E. Aldrich, H. R. Skeggs, L. D. Wright, and K. Folkers [8, 9] assigned the structures β, δ-dihydroxy-β-methylvaleric acid (I) and β-hydroxy-β-methylvalerolactone (II) to the acid and lactone forms of the factors respectively. The generic name mevalonic acid was subsequently used for the new biologically active factor.

SYNTHESES OF DL-MEVALONIC ACID

This structural conclusion was verified by an unequivocal synthesis [9]. The successive stages (III, IV, V, II) are shown in figure 1. The microbiological activity of the synthetic mevalolactone was one half that of the optically active substance obtained from distillers' solubles.

An improved and very useful basic synthesis by C. H. Hoffman, A. F. Wagner, A. N. Wilson, E. Walton, C. H. Shunk, D. E. Wolf, F. W. Holly, and K. Folkers [10] was based on condensation of 4-acetoxy-2-butanone (VI, figure 1) with ethyl bromoacetate. Hydrolysis of the product (VII) gave an equilibrium mixture of the corresponding dihydroxy acid and lactone. The pure lactone was isolated by distillation; the pure acid was obtained as the *bis* N,N′-dibenzylethylenediammonium salt after reacting the acid-lactone mixture with an excess of N,N′-dibenzylethylenediamine in semi-aqueous media.

SYNTHESES OF RADIOACTIVE DL-MEVALONIC ACID

Study of the role of mevalonic acid in polyisoprenoid biosynthesis required labelled mevalonic acid, and the following syntheses were utilized.

J. W. Cornforth and his colleagues [11] synthesized 2-^{14}C-mevalolactone from 4-acetoxy-2-butanone (VI) and ethyl 2-^{14}C-bromoacetate (figure 1). The synthesis of 4-^{14}C-mevalolactone [12] utilized key reactions from both of the earlier syntheses of mevalolactone. The sequence of reactions is shown in figure 2.

An excellent synthesis of 3′,4-^{13}C-mevalolactone was introduced [12] that started from doubly labelled diketene (X). The sequence of reactions is shown in figure 3.

SYNTHESIS OF (+)-MEVALONIC ACID

The biologically active form of the acid is dextrorotatory. It has been obtained by chemical resolution [13], and (−)-mevalonic acid has been obtained by enzymic resolution of the racemate [14]. (+)-Mevalonic acid has been prepared directly by enzymic synthesis [15, 16]; 70 per cent of the (+)-isomer of the racemate was isolated as phospho-(+)-mevalonic acid after selective enzymic phosphorylation with mevalonic kinase in the presence of ATP.

U. Henning, E. M. Möslein, and F. Lynen [16] accomplished the synthesis of (+)-mevalonic acid

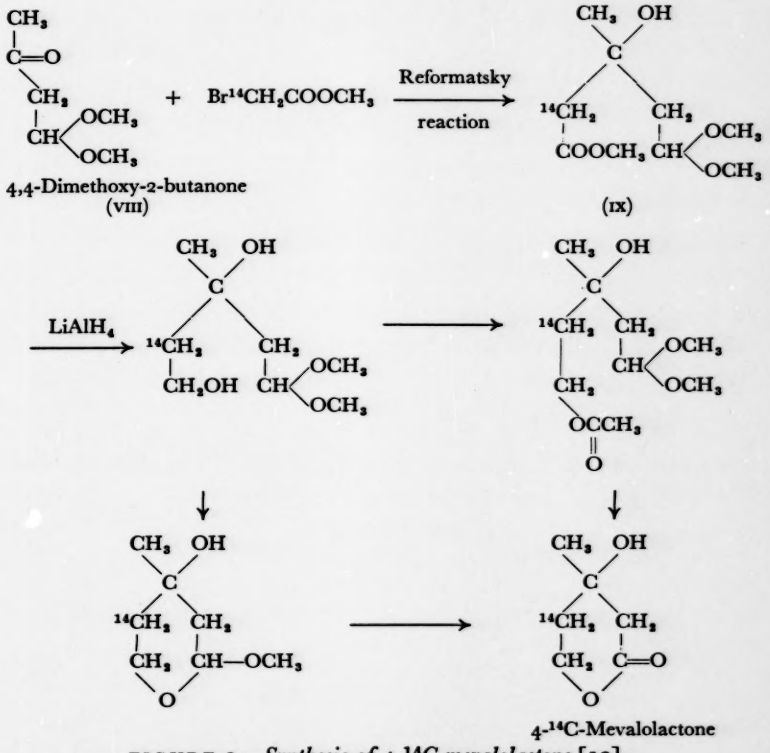


FIGURE 2 – *Synthesis of* 4-^{14}*C-mevalolactone* [12].

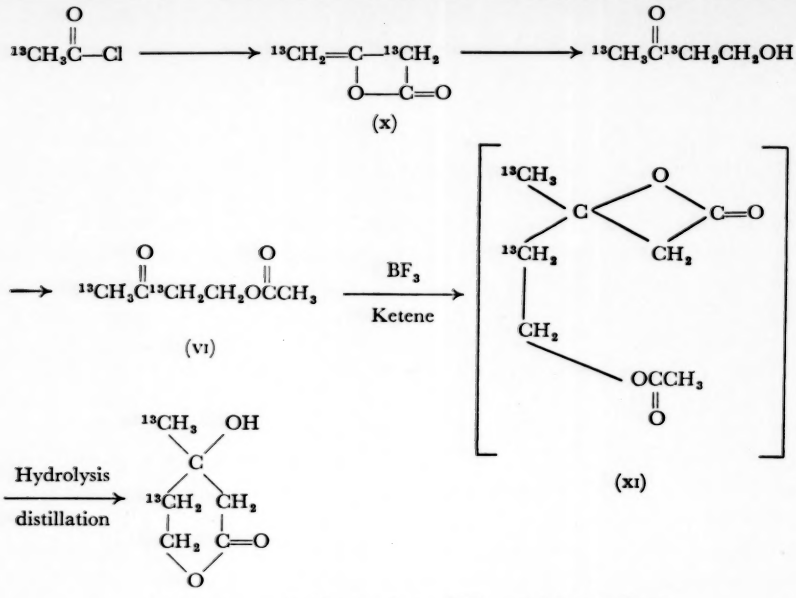


FIGURE 3 – *Synthesis of* 3′,4-^{13}*C-mevalolactone* [12].

from β-methylcrotonyl coenzyme A by the reaction sequence shown below, which was carried out in two stages. The carboxylation reaction was completed as a single step, but the hydration and reduction reactions were combined to prevent cleavage of β-hydroxy-β-methylglutaryl coenzyme A to acetoacetate by an aldolase present in the enzyme system.

β-methylcrotonyl CoA + ATP + CO_2 ⟶
β-methylglutaconyl CoA + ADP + P_i

β-methylglutaconyl CoA + H_2O ⟶
β-hydroxy-β-methylglutaryl CoA

β-hydroxy-β-methylglutaryl CoA + 2TPNH + $2H^+$ ⟶
(+)-mevalonic acid + $2TPN^+$

CONFIGURATION OF MEVALONIC ACID

(+)-Mevalonic acid has been related to (+)-glyceraldehyde by way of quinic acid (figure 4, XII). M. Eberle and D. Arigoni [17] converted quinic acid (XII) to the 'unnatural' (+)-form of mevalolactone by the sequence of reactions outlined in figure 4. Since these reactions yielded the biologically inactive *S*-(+)-lactone (XIII), the biologically active forms of the compound are designated *R*-(−)-mevalolactone and *R*-(+)-mevalonic acid. Additional confirmation for this conclusion was obtained from a comparison of the optical-rotatory dispersion curve of the benzhydrylamide derived from the *S*-(+)-lactone (XIII) with that reported for the benzhydrylamide derived from he antipode of mevalonic acid.

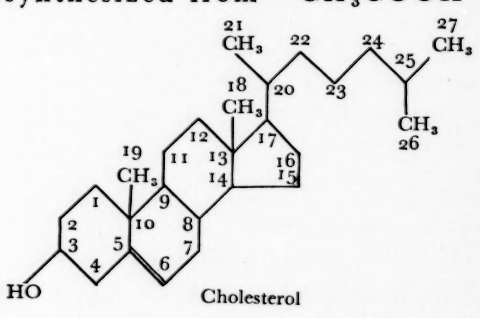
MEVALONIC ACID IN BIOSYNTHESIS

The biosynthesis of sterols

The biosynthesis of sterols and terpenes has been investigated by biochemists and biologists for the past two decades. Chronologically, significant landmarks in following the biosynthetic route (figure 5) were the discoveries that acetate, squalene, lanosterol, mevalonic acid, and isopentenyl pyrophosphate were effective precursors of cholesterol.

Acetate

The role of acetate as a precursor for cholesterol was established in animal-feeding experiments, and by using labelled acetate *in vitro*. K. Bloch and D. Rittenberg [18] observed that the ingestion of deutero-acetic acid by mice and rats led to the formation of deuterocholesterol; the distribution of deuterium showed that acetate was a precursor for both the side chain and the nucleus. Other data excluded propionic, butyric, and succinic acids—as well as higher fatty acids, such as palmitic and stearic acids—from direct participation in the biosynthesis. After the addition of $^{13}CH_3{}^{14}COOH$ and $^{14}CH_3{}^{13}COOH$ to rat-liver homogenates, the ratio of ^{13}C to ^{14}C in the biosynthetic cholesterol showed that approximately 20 per cent more methyl groups than carboxyl groups had been incorporated [19]. It was therefore assumed that some decarboxylation took place in the course of biosynthesis. Selective degradation of the radioactive cholesterol confirmed that acetate serves as a precursor for both the nucleus and side chain. J. W. Cornforth, G. D. Hunter, and G. Popják [20] determined the distribution of radioactive carbon in cholesterol biosynthesized from $^{14}CH_3COOH$ and



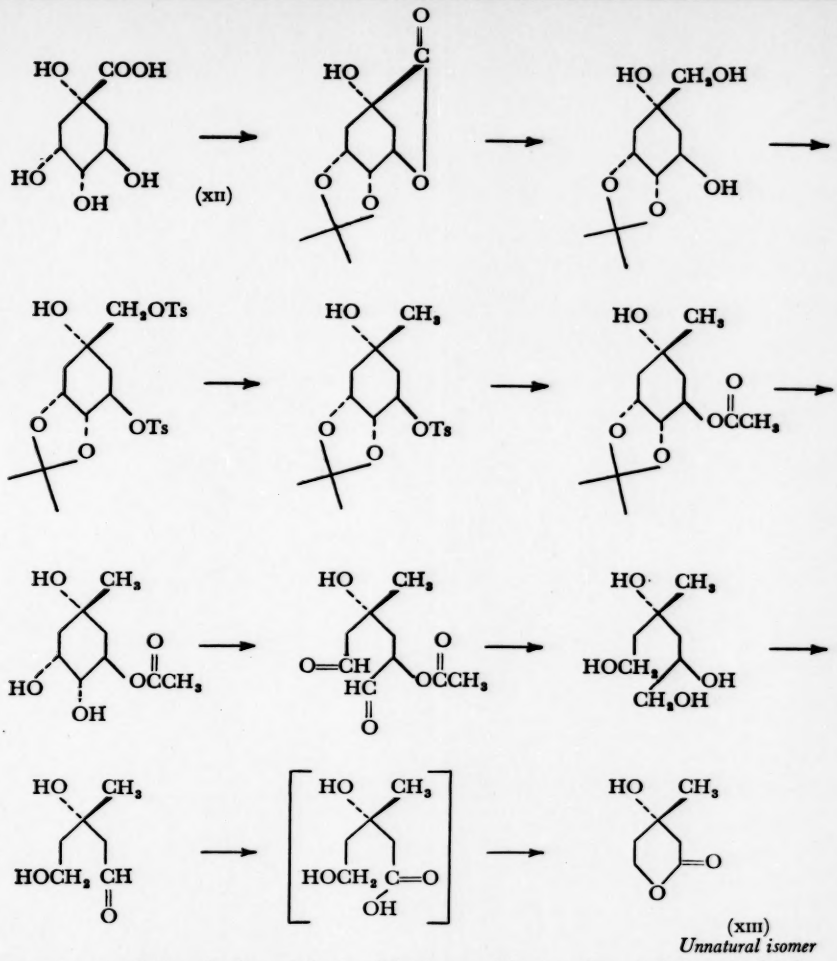


FIGURE 4 – *Stereoselective synthesis of* (+)*-mevalolactone from quinic acid* [17]. (*Ts = toluene sulphonate*).

$CH_3{}^{14}COOH$ by rat-liver slices; fifteen carbon atoms of the sterol were derived from the methyl group and twelve from the carboxyl group of acetate. The same conclusion had been reached by H. N. Little and K. Bloch [19].

Selective degradation of the biosynthetic cholesterol showed that carbon atoms 2, 4, 6, 8, and 10 were derived from the carboxyl group, and carbon atoms 1, 3, 5, and 19 were derived from the methyl group of acetate [20]. J. Wuersch, R. L. Huang, and K. Bloch [21] showed that the entire side chain of the sterol molecule was derived from acetate. After the addition of labelled acetate to an *in vitro* system and selective degradations of the side-chain, it was revealed that carbon atoms 21, 22, 24, 26, and 27 were derived from the methyl group of acetate, and that carbon atoms 20, 23, and 25 were derived from the carboxyl group. When these investigators considered the biosynthetic derivation of the carbon atoms of the side-chain on the basis of the Bonner and Arreguin mechanism [22] (figure 6) for the conversion of acetate to isoprene, it was concluded that five-carbon units related to isoprene participate in the biosynthesis of cholesterol from acetate.

The reaction scheme shown in figure 6 received further support when it was demonstrated by R. O. Brady and S. Gurin [23] that methyl-labelled acetoacetate was incorporated into cholesterol by rat-liver slices without degradation to two-carbon units. Since acetoacetate functions as a cholesterol precursor, it was suggested [24] that the next step of biosynthesis might involve condensation of acetoacetyl coenzyme A with acetyl coenzyme A to yield β-hydroxy-β-methylglutaryl coenzyme A (XIV). They demonstrated that DPN, ATP, and coenzyme A were required for the incorporation of labelled acetate, pyruvate, and acetyl coenzyme A into cholesterol by particle-free aqueous extracts of rat liver, and also observed the conversion of labelled acetate into β-hydroxy-β-methylglutarate in the same extract. H. Rudney [25] showed that radioactivity from labelled acetate was incorporated into β-hydroxy-β-methylglutaric and β,β-dimethylacrylic acids by liver homogenates and extracts. The distribution of radioactivity in the carbon skeleton of each compound was established by degradation and was compatible with the idea that such units could be utilized in biosynthesis without first being degraded to two-carbon units. A detailed investigation of the reactions leading to the conversion of acetate to hydroxymethylglutarate (XIV) conclusively demonstrated that β-hydroxy-β-methylglutaryl coenzyme A is formed from the enzymic condensation of acetyl coenzyme A and acetoacetyl coenzyme A with

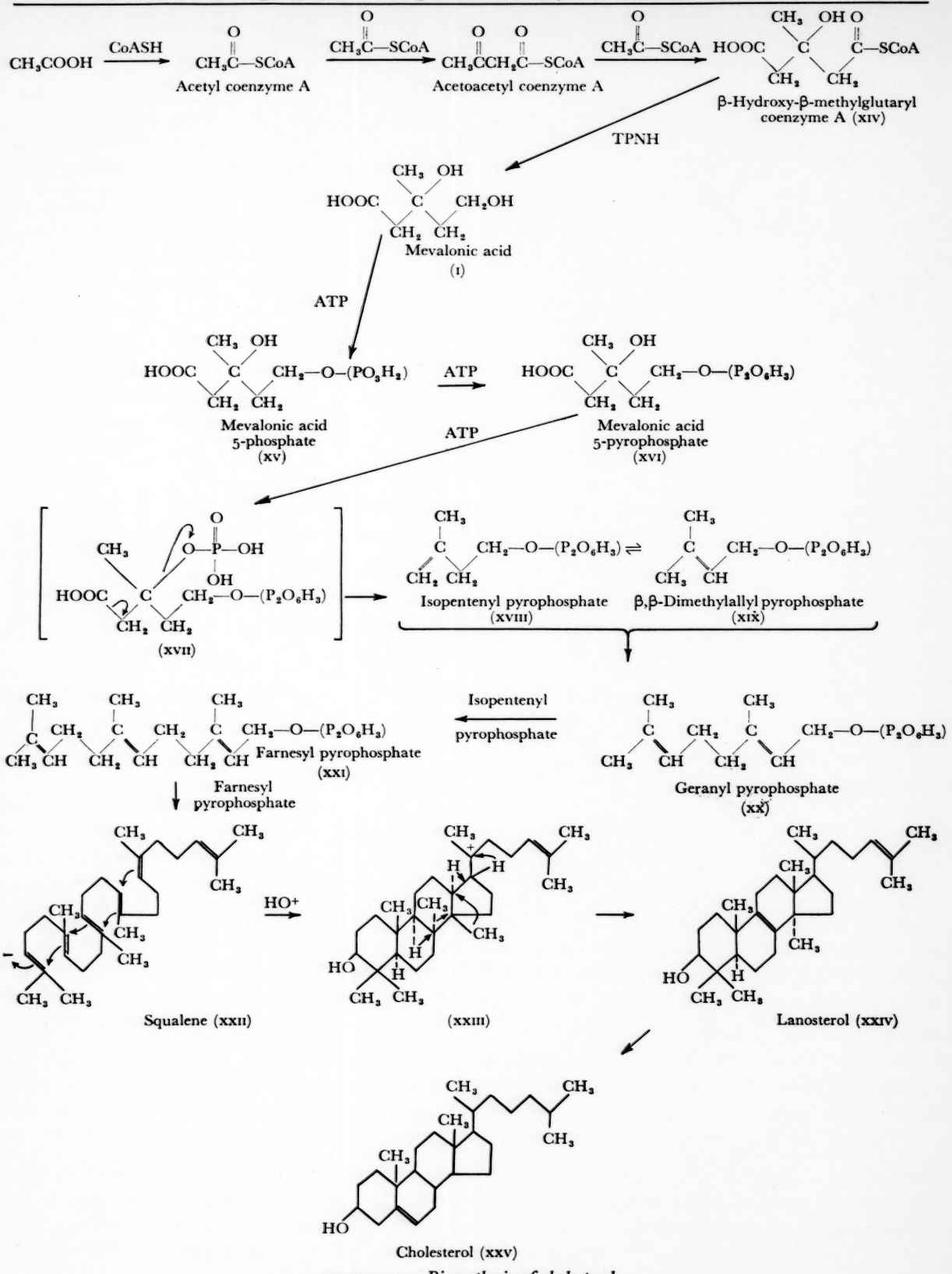


FIGURE 5 – *Biosynthesis of cholesterol.*

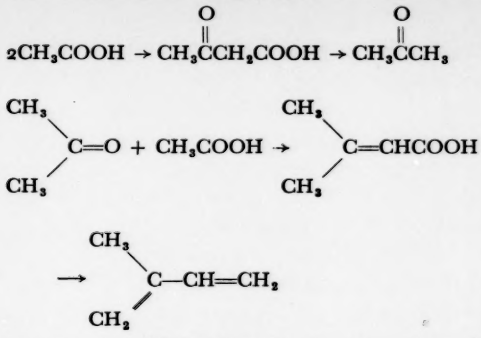


FIGURE 6 – *Mechanism for the conversion of acetate to isoprene* [22].

the concomitant hydrolysis of a thioester bond [26].

Mevalonate

Knowledge of the microbiological activity of mevalonic acid, and of the structural similarity between mevalonic acid and β-hydroxy-β-methylglutaric acid prompted experiments on the possible role of the new factor in cholesterol biosynthesis. Using cell-free rat liver homogenates, P. A. Tavormina, M. H. Gibbs, and J. W. Huff [27] showed by radioactivity experiments that 43·4 per cent of DL-2-^{14}C-mevalonic acid was incorporated into cholesterol; by contrast, only 0·16 per cent and 3·8 per cent respectively of the isotope from 3′-^{14}C-β-hydroxy-β-methylglutaric acid and 4-^{14}C-β, β-dimethylacrylic acid was incorporated into the cholesterol under comparable conditions. Assuming only one of the optical antipodes of the racemate to be utilized, it was apparent that mevalonic acid was very efficiently incorporated. It was also clear that mevalonic acid is an important biological precursor for steroids and for isoprenoid natural products in general, and that it is closely related to 'active isoprene', which is discussed later. Further studies [28] of the incorporation of ^{14}C isotope from 2-^{14}C-mevalonic acid and 1-^{14}C-mevalonic acid showed that the carboxyl group of mevalonic acid is eliminated during the biosynthesis of cholesterol; all the remaining carbon atoms are incorporated.

Mevaldic acid

Since one carboxyl group of β-hydroxy-β-methylglutaric acid is reduced to a hydroxymethyl group during the enzymic synthesis of mevalonic acid, the possible intermediate aldehydic derivative, mevaldic acid, has been studied. It was synthesized [13, 29] from the dimethyl acetal of acetoacetaldehyde and ethyl bromoacetate in a manner analogous to a synthesis of mevalonic acid (figure 1). Mevaldic acid is a very reactive aldehyde, whose presence in solution was readily established by reduction to mevalonic acid and in other ways. Although the microbiological activity of DL-mevaldic acid for *L. acidophilus* is only 0·5 per cent of that of DL-mevalonic acid, the efficiency with which it is incorporated into cholesterol by rat liver is indistinguishable from that of mevalonic acid [30].

Although mevaldic acid has not been detected as an intermediate in the enzymic reduction of β-hydroxy-β-methylglutarate to mevalonate, it has been demonstrated that it is reduced to mevalonic acid by TPNH and yeast enzymes.

Squalene

Mevalonic acid is also related to squalene (XXII), which R. G. Langdon and K. Bloch [31] showed to be a precursor of cholesterol in the intact animal; the utilization of squalene in the rat is ten to twenty times more efficient than that of acetate. J. W. Cornforth and G. Popják [32] studied the biosynthesis of squalene from labelled acetate *in vitro*, and elucidated the derivation of the carbon atoms in the squalene skeleton. The pattern of labelling in squalene was consistent with this being a direct precursor, because proper folding of labelled squalene could yield a sterol nucleus labelled in just the same way as has been established experimentally. These experiments substantially confirmed a biosynthetic pathway from acetate to sterol through squalene.

The first knowledge about intermediates between mevalonic acid and squalene was gained from studies with yeast extracts. From these investigations, mevalonic acid 5-phosphate (XV), mevalonic acid 5-pyrophosphate (XVI), 3-methyl-3-butenyl 1-pyrophosphate (XVIII, isopentenyl pyrophosphate), geranyl pyrophosphate (XX), and farnesyl pyrophosphate (XXI) have been identified as successive intermediates in the biosynthesis of squalene (figure 5).

Because ATP can be shown to be necessary for the enzymic reaction, it was suggested [33] that phosphorylated intermediates are formed in the conversion of mevalonic acid to squalene. T. T. Tchen [34] showed that the enzyme in yeast extract that converts mevalonic acid to squalene is a kinase; in the presence of ATP, this converts mevalonic acid to mevalonic acid 5-phosphate (XV). The latter was converted to squalene in the enzyme system, and it was concluded that mevalonic acid 5-phosphate is the first intermediate in

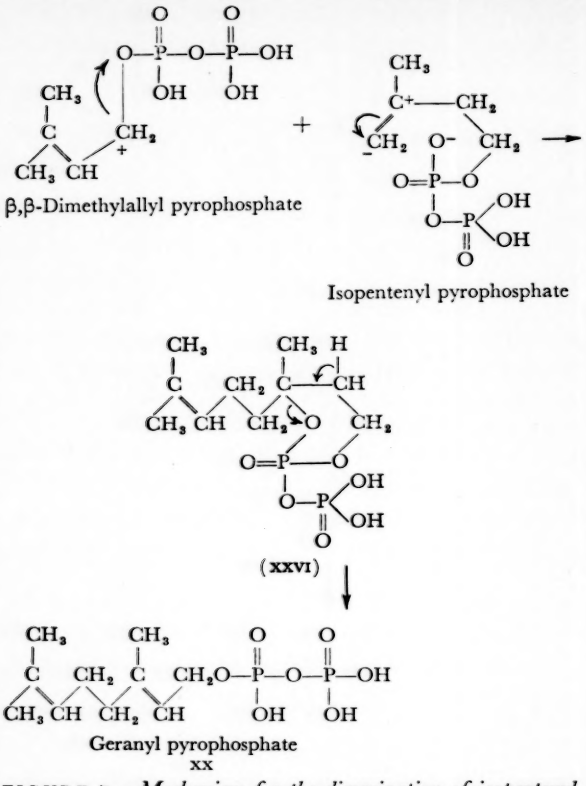


FIGURE 7 – *Mechanism for the dimerization of isopentenyl pyrophosphate to geranyl pyrophosphate* [38, 40–42].

the pathway. The existence of other phosphorylated intermediates of mevalonic acid was predicted [35] on the basis that ATP is required to convert mevalonic acid 5-phosphate to squalene. One of the new phosphorylated derivatives detected in this investigation was converted to squalene in the absence of ATP.

Two products from the enzymic conversion of mevalonic acid 5-phosphate by yeast in the presence of ATP have been isolated and identified [36, 37]. The first was identified as mevalonic acid 5-pyrophosphate (XVI). Radioactivity could be incorporated into this intermediate if either 1-^{14}C- or 2-^{14}C-mevalonic acid 5-phosphate was the starting material; it follows that conversion to the diphospho intermediate (XVI) proceeded with retention of the carboxyl group [37]. The second intermediate was identified as 3-methyl-3-butenyl 1-pyrophosphate (XVIII). As the formation of isopentenyl pyrophosphate from mevalonic acid 5-pyrophosphate requires ATP, phosphorylation must be an essential part of the elimination reaction. For the conversion of mevalonic acid 5-pyrophosphate to isopentenyl pyrophosphate, it was assumed [37] that phosphorylation occurs at the tertiary hydroxyl group, since the chemical lability of a tertiary ester would be expected to lead to unsaturation. The mechanism for the reaction leading from the triphospho intermediate (XVII) to isopentenyl pyrophosphate is visualized as a concerted decarboxylation-dephosphorylation [37]; protonation of the carbon chain was not included as a step in the mechanism, because earlier studies had shown that deuterium was not introduced during the biosynthesis of squalene in deuterium oxide.

The enzymic isomerization of isopentenyl pyrophosphate (XVIII) to β,β-dimethylallyl pyrophosphate (XIX) has been observed in aqueous extracts of yeast by F. Lynen and his colleagues [38, 39]. Geranyl pyrophosphate (XX) and farnesyl pyrophosphate (XXI) were also identified by these investigators as intermediates in the biosynthesis of squalene from mevalonic acid by cell-free yeast extracts. A mechanism, suggested by Bloch, Cornforth, and Lynen, for the synthesis of geranyl pyrophosphate from isopentenyl pyrophosphate is visualized (figure 7) as the alkylation of an isolated double bond by an allylic carbonium ion [38, 40, 41]. Isopentenyl pyrophosphate isomerizes to β,β-dimethylallyl pyrophosphate and probably exists as an equilibrium mixture of the two forms. One of the isomers can be regarded as the electrophile and the other as the nucleophile necessary for the condensation of two molecules to yield a dimer. The unsaturated moiety of isopentenyl

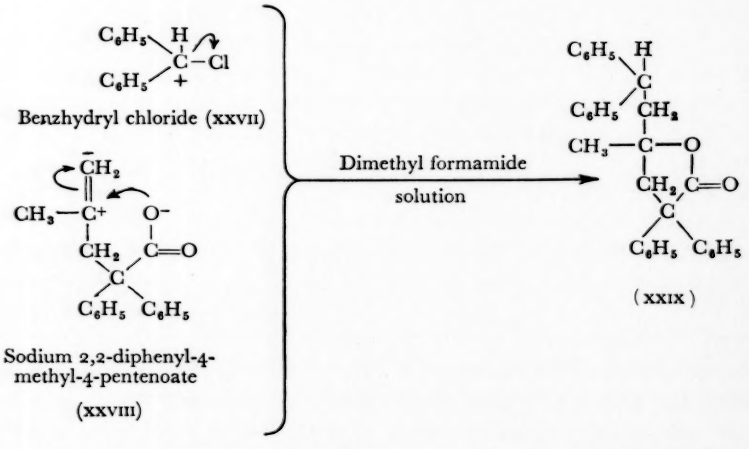


FIGURE 8 – *An alkylation and cyclization reaction analogous to the dimerization of 'active isoprene'* [42].

pyrophosphate provides the nucleophilic centre; β,β-dimethylallyl pyrophosphate, on the basis of analogy to the behaviour of benzyl and other allyl esters, is susceptible to attack at the oxygenated carbon atom by the nucleophile.

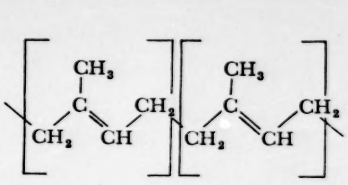
W. S. Johnson and R. A. Bell [42] have suggested that the alkylation reaction proceeds with the anchimeric assistance of the pyrophosphate moiety of isopentenyl pyrophosphate, leading to the cyclic intermediate (XXVI). The cyclic tertiary phosphate moiety should be highly unstable and should collapse readily to yield geranyl pyrophosphate. Support for the existence of such an intermediate was derived from a study of the alkylation reaction shown in figure 8. Isolation of the γ-lactone (XXIX) of the predictable alkylation product not only supports the concept of a cyclic tertiary phosphate being formed during the polymerization of isopentenyl pyrophosphate, but is the basis for a new approach to cyclization reactions under neutral, or even alkaline, conditions.

The biosynthetic pathway from mevalonic acid to squalene in rat liver is similar to that in yeast cells (figure 5). According to G. Popják and his colleagues [43–46], the sequence includes such reactions as the conversion of mevalonic acid to mevalonic acid 5-phosphate, mevalonic acid 5-pyrophosphate, and isopentenyl pyrophosphate (XVIII). The sequence continues with the isomerization of the last of these to dimethylallyl pyrophosphate (XIX) and the condensation of these isomers to yield geranyl pyrophosphate (XX), which condenses with isopentenyl pyrophosphate to yield farnesyl pyrophosphate.

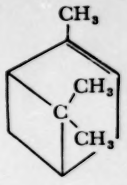
Lanosterol

Once it had been established that squalene was an intermediate in the biogenesis of cholesterol, R. B. Woodward and K. Bloch [47] proposed that lanosterol (XXIV) was intermediate between squalene and cholesterol on the biosynthetic pathway. The biosynthesis of lanosterol from labelled acetate in a tissue homogenate was demonstrated [48], and the further conversion of this labelled lanosterol to cholesterol was shown.

The cyclization of squalene to lanosterol is regarded as a single reaction involving the concerted participation of at least sixteen centres. A general mechanism proposed by A. Eschenmoser, L. Ruzicka, O. Jeger, and D. Arigoni [49] visualizes the formation of an intermediate carbonium ion after the reaction of a cationic activator, such as HO^+, with a properly folded substrate; for the conversion of squalene to lanosterol the reaction may be visualized as a series of nucleophilic attacks on electron-deficient centres by the electrons of double bonds. With a suitably folded squalene molecule, the concerted shift of electrons initiated by HO^+ is transmitted through five double bonds to form the tetracyclic structure (XXIII) containing a carbonium ion at C-20. After the 1,2-migration of two methyl groups and two hydrogen atoms, a proton is eliminated from the 9-position to form the stable lanosterol molecule.



Rubber
(XXX)



α-Pinene
(XXXI)

Cholesterol

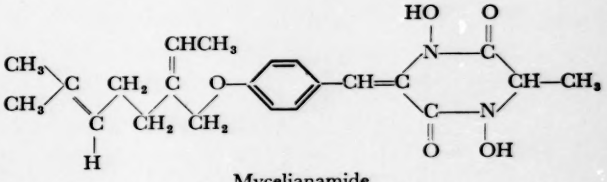
The conversion of lanosterol to cholesterol requires the oxidative cleavage of three methyl groups. Several intermediates and new precursors which function between lanosterol and cholesterol have been identified. Detailed reviews of this research have been published by J. W. Cornforth [50] and P. Crabbé [51].

'ACTIVE ISOPRENE'

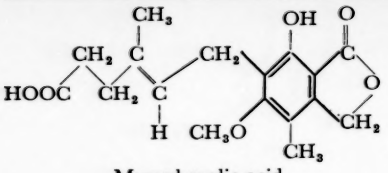
A critical step in the participation of mevalonic acid in cholesterol biosynthesis is its conversion to isopentenyl pyrophosphate (XVIII), which is regarded as 'active isoprene'. This compound, with its isomer β,β-dimethylallyl pyrophosphate, constitutes the basic unit for biochemical polymerization. When polymers arising from this basic unit are subjected to other biochemical transformations, a wide variety of naturally occurring isoprenoid compounds and polymers result.

THE BIOSYNTHESIS OF ISOPRENOIDS

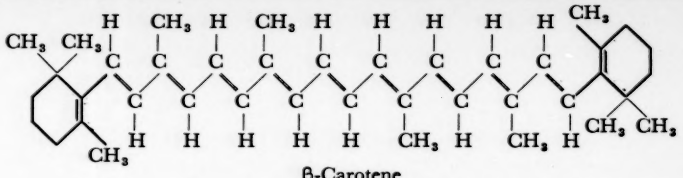
The importance of mevalonic acid in the biosynthesis of isoprenoid natural products has been substantiated by numerous investigations, based



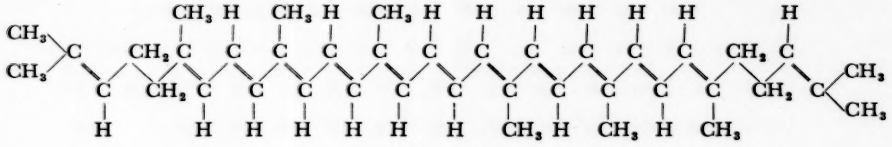
Mycelianamide
(XXXII)



Mycophenolic acid
(xxxiii)



β-Carotene
(xxxiv)

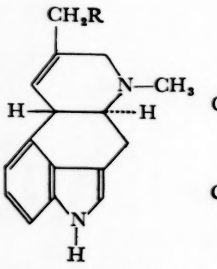


Lycopene
(xxxv)

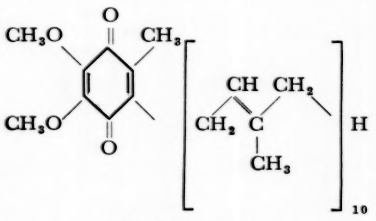
on the incorporation of radioactive mevalonic acid into a variety of molecules, such as rubber, terpenes, and alkaloids.

R. B. Park and J. Bonner [52] observed the incorporation of 2-^{14}C-mevalonic acid into rubber (xxx) by a crude enzyme system from latex of *Hevea brasiliensis*; the efficiency of conversion was about 2 per cent. Subsequent studies [53] showed a 9·3 per cent conversion of mevalonic acid to polyisoprene. The conversion of isopentenyl pyrophosphate into rubber has been demonstrated [54].

In the terpene field, R. G. Stanley [55] showed that seedlings of *Pinus attenuata* incorporate 2-^{14}C-mevalonic acid into the monoterpene fraction, which is composed almost entirely of α-pinene (xxxi). The incorporation of 2-^{14}C-mevalonic acid into the ten-carbon terpenoid side chain of mycelianamide (xxxii) by *Penicillium griseofulvum* and into the seven-carbon side chain of mycophenolic acid (xxxiii) by *Penicillium brevi-compactum* was also demonstrated [56]. The seven-carbon unit is regarded as being derived from a geranyl moiety.



Agroclavine (xxxvi, R = H) and elymoclavine (xxxvii, R = OH)



Coenzyme Q_{10}
(xxxviii)

The incorporation of 2-^{14}C-mevalonic acid into β-carotene (xxxiv) by growing cultures of the fungus *Phycomcyes blakesleeanus* has been demonstrated by G. D. Braithwaite and T. W. Goodwin [57], who also observed incorporation of the labelled factor into β-carotene by isolated slices of carrot root. It has been shown [58] that the mould *Mucor hiemalis* can incorporate 2-^{14}C-mevalonic acid into β-carotene. Enzymic synthesis of lycopene (xxxv) from 2-^{14}C-mevalonic acid in tomato homogenates has been observed [59].

The biosynthetic significance of mevalonic acid has been demonstrated in other fields too. Thus it has been shown [60] that *Claviceps purpurea*, the parasitic fungus of rye, incorporates ^{14}C isotope from 2-^{14}C-mevalolactone into the alkaloids agroclavine (xxxvi) and elymoclavine (xxxvii). This, and 4-^{14}C-mevalonic acid, can be incorporated into the isoprenoid side chain of coenzyme Q_{10} (xxxviii) [61, 62].

Figures 2, 3, and 4 are reproduced from 'Advances in Enzymology', Volume XXIII, by permission of Interscience Publishers Inc.

REFERENCES

[1] Novak, A. F., Hauge, S. M., and Carrick, C. W. *Poult. Sci.*, **26**, 604, 1947.

[2] Novak, A. F. and Hauge, S. M. *J. biol. Chem.*, **174**, 647, 1948.

[3] Skeggs, H. R., Wright, L. D., Cresson, E. L., MacRae, G. D. E., Hoffman, C. H., Wolf, D. E., and Folkers, K. *J. Bact.*, **72**, 519, 1956.

[4] Ott, W. H., Dickinson, A. M., van Inderstine, A., Bazemore, A. W., Page, A. C., and Folkers, K. *J. Nutr.*, **64**, 525, 1958.

[5] Tamura, G. *J. gen. appl. Microbiol.*, **2**, 431, 1956.

[6] Tamura, G. and Folkers, K. *J. org. Chem.*, **23**, 772, 1958.

[7] Wright, L. D., Cresson, E. L., Skeggs, H. R., MacRae, G. D. E., Hoffman, C. H., Wolf, D. E., and Folkers, K. *J. Amer. chem. Soc.*, **78**, 5273, 1956.

[8] Wolf, D. E., Hoffman, C. H., Aldrich, P. E., Skeggs, H. R., Wright, L. D., and Folkers, K. *Ibid.*, **78**, 4499, 1956.

[9] *Idem. Ibid.*, **79**, 1486, 1957.
[10] HOFFMAN, C. H., WAGNER, A. F., WILSON, A. N., WALTON, E., SHUNK, C. H., WOLF, D. E., HOLLY, F. W., and FOLKERS, K. *Ibid.*, **79**, 2316, 1957.
[11] CORNFORTH, J. W., CORNFORTH, R. H., POPJÁK, G., and YOUHOTSKY GORE, I. *Biochem. J.*, **69**, 146, 1948.
[12] CORNFORTH, J. W., CORNFORTH, R. H., PELTER, A., HORNING, M. G., and POPJÁK, G. *Tetrahedron*, **5**, 311, 1959.
[13] SHUNK, C. H., LINN, B. O., HUFF, J. W., GILFILLAN, J. L., SKEGGS, H. R., and FOLKERS, K. *J. Amer. chem. Soc.*, **79**, 3294, 1957.
[14] LYNEN, F. and GRASSL, M. *Z. physiol. Chem.*, **313**, 291, 1958.
[15] LYNEN, F., EGGERER, H., HENNING, U., and KESSEL, I. *Angew. Chem.*, **70**, 738, 1958.
[16] HENNING, U., MÖSLEIN, E. M., and LYNEN, F. *Arch. Biochem. Biophys.*, **83**, 259, 1959.
[17] EBERLE, M. and ARIGONI, D. *Helv. Chim. Acta*, **43**, 1508, 1960.
[18] BLOCH, K. and RITTENBERG, D. *J. biol. chem.*, **143**, 297, 1942; **145**, 625, 1942.
[19] LITTLE, H. N. and BLOCH, K. *J. biol. Chem.*, **183**, 33, 1950.
[20] CORNFORTH, J. W., HUNTER, G. D., and POPJÁK, G. *Biochem. J.*, **54**, 597, 1953; *Arch. Biochem. Biophys.*, **42**, 482, 1953.
[21] WUERSCH, J., HUANG, R. L., and BLOCH, K. *J. biol. Chem.*, **195**, 439, 1952.
[22] BONNER, J. and ARREGUIN, B. *Arch. Biochem. Biophys.*, **21**, 109, 1949.
[23] BRADY, R. O. and GURIN, S. *J. biol. Chem.*, **189**, 371, 1951.
[24] RABINOWITZ, J. L. and GURIN, S. *Ibid.*, **208**, 307, 1954.
[25] RUDNEY, H. *J. Amer. chem. Soc.*, **76**, 2595, 1954; **77**, 1698, 1955.
[26] *Idem. J. biol. Chem.*, **227**, 363, 1957.
[27] TAVORMINA, P. A., GIBBS, M. H., and HUFF, J. W. *J. Amer. chem. Soc.*, **78**, 4498, 1956.
[28] TAVORMINA, P. A. and GIBBS, M. H. *Ibid.*, **78**, 6210, 1956.
[29] EGGERER, H. and LYNEN, F. *Liebigs Ann.*, **608**, 71, 1956.
[30] FOLKERS, K., SHUNK, C. H., LINN, B. O., ROBINSON, F. M., WITTREICH, P. E., HUFF, J. W., GILFILLAN, J. L., and SKEGGS, H. R. 'Ciba Foundation Symposium on Biosynthesis of Terpenes and Sterols', p. 20. J. and A. Churchill Ltd, London. 1959.
[31] LANGDON, R. G. and BLOCH, K. *J. Amer. chem. Soc.*, **74**, 1869, 1952.
[32] CORNFORTH, J. W. and POPJÁK, G. *Biochem. J.*, **58**, 403, 1954.
[33] ADMUR, B. H., RILLING, H., and BLOCH, K. *J. Amer. chem. Soc.*, **79**, 2646, 1957.
[34] TCHEN, T. T. *Ibid.*, **79**, 6344, 1957; *J. biol. Chem.*, **233**, 1100, 1958.
[35] PHILLIPS, A. H., TCHEN, T. T., and BLOCH, K. *Fed. Proc.*, **17**, 289, 1958.
[36] CHAYKIN, S., LAW, J., PHILLIPS, A. H., TCHEN, T. T., and BLOCH, K. *Proc. Nat. Acad. Sci.*, **44**, 998, 1958.
[37] BLOCH, K., CHAYKIN, S., PHILLIPS, A. H., and DE WAARD, A. *J. biol. Chem.*, **234**, 2595, 1959.
[38] LYNEN, F., EGGERER, H., HENNING, U., and KESSEL, I. *Angew. Chem.*, **70**, 738, 1958.
[39] LYNEN, F., AGRANOFF, B. W., EGGERER, H., HENNING, U., and MÖSLEIN, E. M. *Angew. Chem.*, **71**, 657, 1959.
[40] RILLING, H. C. and BLOCH, K. *J. biol. Chem.*, **234**, 1424, 1959.
[41] CORNFORTH, J. W. *Tetrahedron Lett.*, No. 19, 29, 1959.
[42] JOHNSON, W. S. and BELL, R. A. *Ibid.*, No. 12, 27, 1960.
[43] DE WAARD, A. and POPJÁK, G. *Biochem. J.*, **73**, 410, 1959.
[44] LEVY, H. R. and POPJÁK, G. *Ibid.*, **72**, 35P, 1959; **75**, 417, 1960.
[45] POPJÁK, G. *Tetrahedron Lett.*, No. 19, 19, 1959.
[46] GOODMAN, DE W. S. and POPJÁK, G. *Biochem. J.*, **74**, 35P, 1960; *J. Lipid Res.*, **1**, 286, 1960.
[47] WOODWARD, R. B. and BLOCH, K. *J. Amer. chem. Soc.*, **75**, 2023, 1953.
[48] CLAYTON, R. B. and BLOCH, K. *J. biol. Chem.*, **218**, 305, 319, 1956.
[49] ESCHENMOSER, A., RUZICKA, L., JEGER, O., and ARIGONI, D. *Helv. Chim. Acta*, **38**, 1890, 1955.
[50] CORNFORTH, J. W. *J. Lipid Res.*, **1**, 3, 1959.
[51] CRABBÉ, P. *Rec. chem. Progr.*, **20**, 189, 1959.
[52] PARK, R. B. and BONNER, J. *J. biol. Chem.*, **233**, 340, 1958.
[53] KEKWICK, R. G. O., ARCHER, B. L., BARNARD, D., HIGGINS, G. M. C., MC SWEENEY, G. P., and MOORE, C. G. *Nature, Lond.*, **184**, 268, 1959.
[54] HENNING, U., MÖSLEIN, E. M., and LYNEN, F. *Biochem. Z.*, **333**, 534, 1961.
[55] STANLEY, R. G. *Nature, Lond.*, **182**, 738, 1958.
[56] BIRCH, A. J., ENGLISH, R. J., MASSY-WESTROPP, R. A., and SMITH, H. *J. chem. Soc.*, 369, 1958.
[57] BRAITHWAITE, G. D. and GOODWIN, T. W. *Biochem. J.*, **66**, 13P, 31P, 1957.
[58] GROB, E. C. *Chimia*, **11**, 338, 1957.
[59] SHNEOUR, E. and ZABIN, I. *J. biol. Chem.*, **234**, 770, 1959.
[60] BIRCH, A. J., MCLOUGHLIN, B. J., and SMITH, H. *Tetrahedron Lett.*, No. 7, 1, 1960.
[61] GLOOR, U. and WISS, O. *Experientia*, **14**, 410, 1958.
[62] GLOOR, U., SCHINDLER, O., and WISS, O. *Helv. Chim. Acta*, **43**, 2089, 1960.

The Mohole

SIR EDWARD BULLARD

It has long been known that there is a change in the properties of the rocks of the earth at a depth of about 35 km under the continents. More recently it has been shown that this discontinuity, called the Moho, lies at a depth of only 5 km below the ocean floor. This article describes the plans to reach this discontinuity by drilling a hole, known as the Mohole, and discusses the results that have been obtained by drilling a trial hole to a lesser depth.

THE MOHOROVIČIĆ DISCONTINUITY

On the morning of the 8th October 1909, an earthquake occurred about 40 km south of Zagreb, in Croatia. Readable records of this earthquake were obtained at stations all over Europe and were examined by Dr Andrija Mohorovičić, the director of the Meteorological Observatory at Zagreb [5]. He found that, as with other earthquakes, there were two kinds of waves present, which could be interpreted as the compressional and distortional waves required by the theory of the propagation of elastic waves through a solid. In compressional or P waves, the oscillatory motion of the particles takes place along a line in the direction in which the wave is travelling, as is shown in figure 1a. In distortional or S waves, the motion of the particles is at right angles to the direction of travel of the wave (figure 1b). The

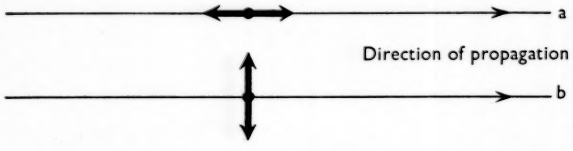


FIGURE 1 – *Motion of a particle in* (a) *a compressional wave and* (b) *a distortional wave.*

existence and significance of these two types of wave had been known for some years; the new discovery made by Mohorovičić was that there are two P waves. At short distances, the first wave to arrive travels at a speed that varies from place to place within the range $5\frac{1}{2}$–$6\frac{1}{2}$ km/s. At about 170 km this is overtaken by another P wave travelling at 8·1 km/s. Over a range of distance on each side of this, usually out to about 800 km, both waves can be recorded, as is shown in the seismogram of figure 5. Beyond this the slower wave fades out. The travel times of the two waves are shown as a function of distance in figure 2.

Mohorovičić suggested that the slower of the two P waves travels by the direct route between

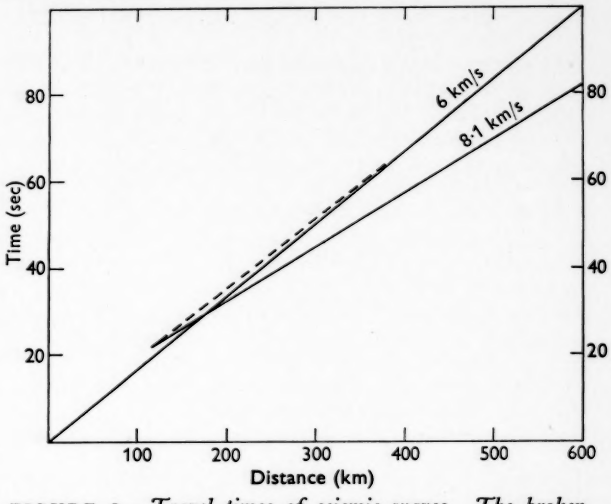


FIGURE 2 – *Travel times of seismic waves. The broken line gives the travel time of the reflected wave.*

the focus of the earthquake and the seismograph, and that the faster wave is refracted at a discontinuity at a depth of about 50 km, as is shown in figure 3. This suggestion has been very widely accepted, and the discontinuity is known as the Mohorovičić discontinuity, or Moho. Mohorovičić overestimated its depth, which varies, from place to place, between 30 and 45 km. If the discontinuity is sharp, it should be possible to observe a wave reflected from the Moho; the path of this wave is shown by a broken line in figure 3. Several workers claim to have detected this reflected wave, but it is difficult to be sure of its existence, because of the continuous motion shown on the record

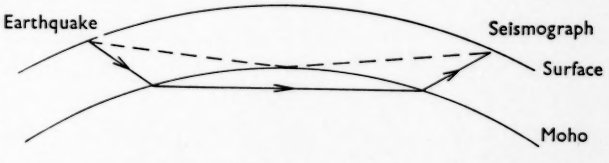


FIGURE 3 – *Section of the Earth, showing the refraction of a seismic wave at the Moho. The broken line shows the track of the reflected wave. The curvature of the Earth is exaggerated.*

FIGURE 4 – *A. Mohorovičić.*

following the P waves. The matter is of some importance, since a reflection is only possible if the change in seismic velocity occurs over a range of depth of less than a wavelength. Unless the reflected waves can be observed it is difficult to be sure whether the Moho is a true discontinuity or whether the change in seismic velocity is spread over a range of depth of a few kilometres.

Observations at sea, using explosions as a source of energy and hydrophones in the water as detectors, have shown a similar discontinuity at a depth of about 5 km beneath the floor of the deep ocean, that is about 10 km beneath the sea surface. It is difficult to trace the Moho through the transition region between a continent and an ocean. The best-investigated region is that off the eastern coast of the United States [3], a generalized section of which is shown in figure 11. There is little doubt that the continental and oceanic Moho are connected in the way shown in the figure. The velocity

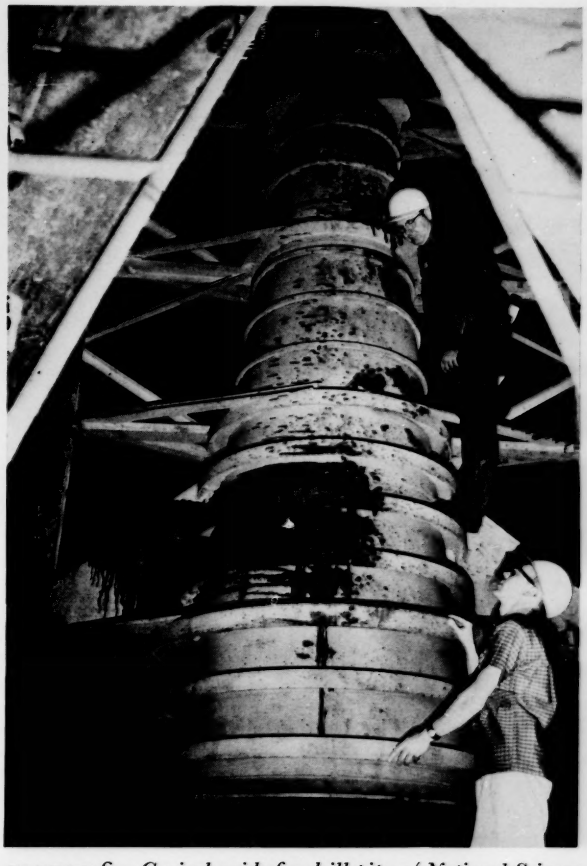


FIGURE 6 – *Conical guide for drill pipe. (National Science Foundation photo.)*

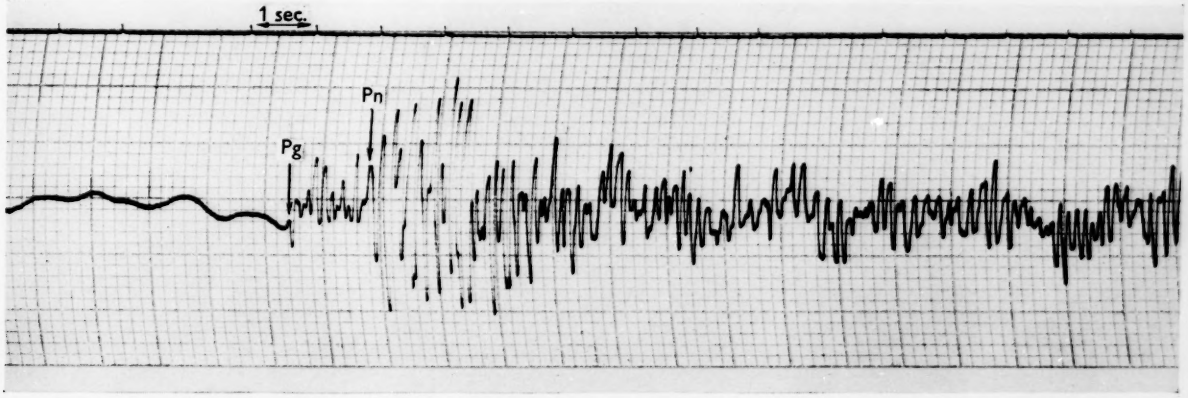


FIGURE 5 – *Seismogram taken at a distance of 113 km from an explosion (approx. 500 kg Torpex) in Loch Striven, near Dunoon, taken by the seismic group of the Atomic Energy Authority. Pg is the arrival of the direct wave and Pn that of the wave refracted at the Moho.*

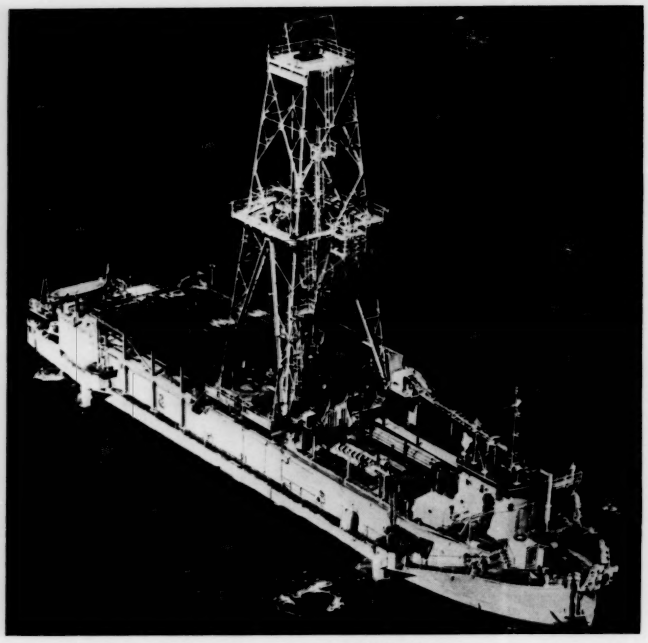


FIGURE 7 – CUSS I.

FIGURE 8 – CUSS I *and float with radar reflector.* (*This figure and figure* 10 *reproduced by courtesy of Global Marine Exploration Company.*)

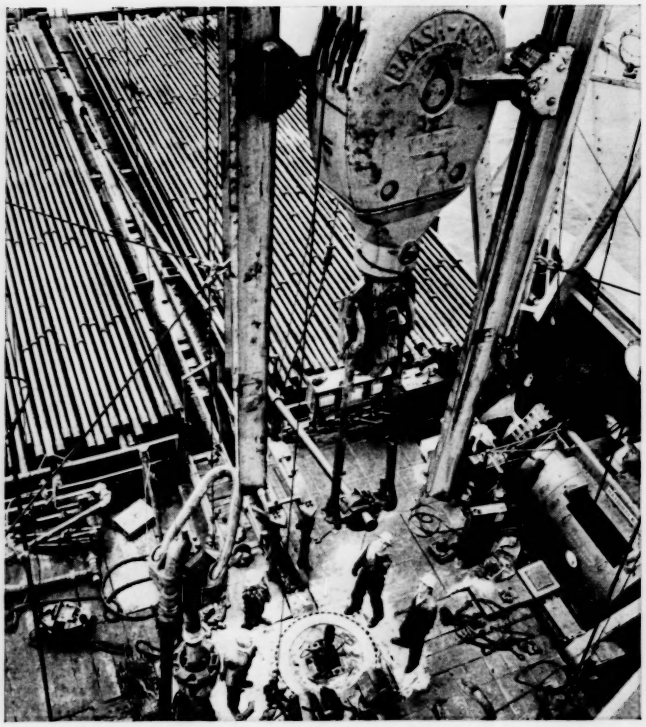


FIGURE 9 – *Raising an* 18 *m length of drill pipe.*

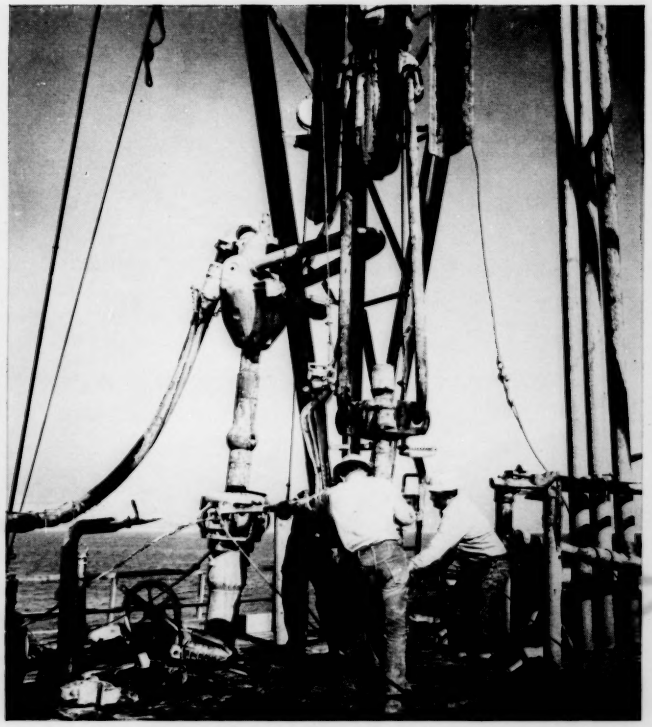


FIGURE 10 – *The rotating table. The drill pipe hangs from the block, the lower part of which is at the top of the figure. The fitting through which sea water is pumped down the drill pipe can be seen on the left.*

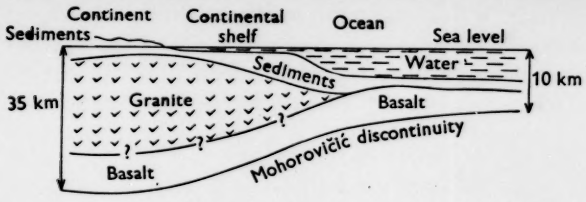


FIGURE 11 – *Section off the eastern coast of the United States.*

below the oceanic discontinuity is usually within 0·2 km/s of the continental average of 8·1 km/s.

Above the discontinuity the continental rocks appear to be different from the oceanic ones. Beneath the sediments of the continents there lie rocks similar to those exposed in the ancient shield areas, such as Canada and east Africa. These rocks are typically granites, gneisses, and schists, and probably represent the worn-down roots of ancient mountains. The seismic velocity above the Moho often increases with depth, probably because of an increase in the proportion of basic rocks. In some places there may be a separate layer of more basic rocks between the granite and the Moho. Beneath the oceans there is a variable thickness of soft sediments in which seismic waves travel a little faster than they do in water. The thickness of this material is normally between 0·1 and 1 km. Beneath this comes 'layer 2' with a seismic velocity in the neighbourhood of 4·8 km/s, and a thickness of about 2 km. The 3 km between this and the Moho is called 'layer 3', with a seismic velocity of $6\frac{1}{2}$–7 km/s. Since there is an enormous preponderance of basalt among oceanic igneous rocks, it is usually supposed that both layer 2 and layer 3 are composed of basalt, perhaps interlayered with consolidated sediments.

The structure shown in figure 11 is based primarily on the seismic results and on petrological probabilities; it is, however, confirmed by measurements of the acceleration due to gravity. On the average, gravity at sea level over a continent is within a few parts in a million of that over the neighbouring ocean in the same latitude. From this it follows that the mass beneath unit area is the same for the continents and the oceans. This suggests that the material beneath the Moho has a higher density than that above, and that the greater amount of this heavy material beneath the ocean compensates for the lightness of the seawater relative to the near-surface rocks of the continents. The density difference required is about 0·5 g/cm^3, which suggests a density of about 3·2 g/cm^3 beneath the Moho. This is an example of the principle of isostasy, which was discovered in the middle of the last century by Archdeacon Pratt of Calcutta, from an examination of the deflections of the vertical observed by the Survey of India. The principle of isostasy states that the Earth's topographic features float in hydrostatic equilibrium in the same way as an iceberg floats in the sea. It is true in a general way for most large features, but applies to averages over considerable areas rather than to individual points. The compensation for the visible masses of topographic irregularities seems to be achieved by variations in the depth of the Moho, and there is a strong correlation of this depth with the height of the topography and with the Bouguer gravity anomaly, as is shown in figure 12. (The Bouguer anomaly represents the attraction or lack of attraction of hidden differences in density after allowance has been made for the visible topography.)

THE MATERIAL BENEATH THE MOHO

The Mohorovičić discontinuity appears to be one of the major features of the earth. It occurs beneath all the continents and all the oceans. Above it there is the infinite variety and complexity of geology; but below it, if we may judge from the constancy of the seismic velocities, there is a much greater uniformity.

In spite of its obvious importance, very little is known about the Moho. The seismic velocities above and below it are known and we have some idea of the density jump across it, but we do not know whether it represents a sharp contact between two different materials or whether the change from one to the other is spread over a range of depth of a few kilometres.

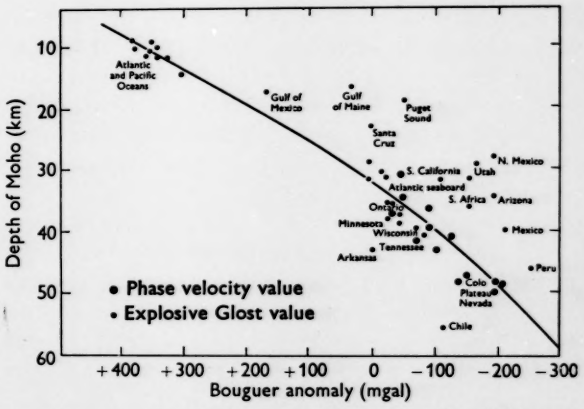


FIGURE 12 – *Correlation between the depth of the Moho and the Bouguer anomaly. The results shown by the large dots were obtained by the method described in the text, those shown by the small dots by a method using surface waves from earthquakes* (*from Woollard* [6]).

Petrology gives some hints as to the nature of the material. There are, in various parts of the world, intrusive rocks containing minerals that can be formed only at high pressures. The best known of these are the diamond pipes of South Africa. The stability relations of diamond and graphite are now well understood, and it is clear that the diamonds must have been formed below the Moho. The associated rocks must, therefore, also come from great depths, and must represent a more or less contaminated sample of the material beneath the Moho. The rocks of the diamond pipes are mineralogically complicated, but consist predominantly of iron and magnesium silicates, particularly olivine, garnet, and enstatite. On the basis of these and other occurrences of rocks derived from great depths it has often been suggested that the material beneath the Moho is either peridotite, which is essentially composed of magnesium-rich olivine, $(Mg,Fe)_2SiO_4$, or eclogite, a rock composed of garnets and pyroxenes.

The comparison of the densities, and elastic wave velocities, of proposed materials with the properties of those actually occurring below the Moho is complicated by the changes in crystal structure found in many silicates at about the pressures and temperatures of the continental Moho. In fact it has been suggested that the discontinuity might be due to phase changes in basalt without any change of material [2, 4]. This is quite a plausible view for the Moho beneath the continents, but can scarcely be correct beneath the oceans. The pressure at the oceanic Moho is only about 1800 kg/cm^2 and the temperature 120° C; a thoroughgoing reorganization of the silicate lattices at so low a temperature and pressure seems improbable. If, as seems likely, the oceanic and continental Mohos are continuous, it is probable that the continental Moho also is not due to a phase change. The expectation is that the Moho is underlain by a material different from that above.

THE MOHOLE PROJECT

Clearly there are many possibilities, and it is unlikely that the nature of the material beneath the Moho can be determined by laboratory experiment or by observations taken at the Earth's surface. The most satisfactory solution of the difficulty would clearly be to drill a hole and obtain samples. The deepest hole so far drilled on land reached a depth of 7·7 km; there seems no prospect of improving techniques sufficiently to drill to the depth of 35 km needed to reach the Moho on the continents. At sea, where the Moho often lies beneath 5 km of water and 5 km of sediments, the prospects appear more hopeful, though there are still formidable difficulties. The present attempt stems from a proposal by H. Hess and W. Munk in 1957 to the American Miscellaneous Society (AMSOC). This society originated in 1952 as an informal and not wholly serious association of American earth scientists, formed 'to encourage miscellaneous projects'. In April 1957, the society formed a deep-drilling committee, with G. Lill as chairman, to consider the possibility of drilling to the Moho. In September of the same year, the triannual meeting of the International Union of Geodesy and Geophysics took place in Toronto. The matter was not on the agenda, but a good deal of private discussion took place and the possibilities received publicity in the press. In spite of considerable scepticism, the Associations of Oceanography and of Seismology were persuaded to adopt resolutions supporting the project.

At this stage it was clear that there was a case good enough to require investigation. As AMSOC had no officers, no defined membership, and no legal existence, its deep-drilling committee was taken over by the National Academy of Sciences, which received a small grant from the National Science Foundation to study ways and means. W. Bascom was appointed director of the project. He recruited the necessary staff and enlisted the help of drilling contractors, oil companies, and naval architects. The hole was christened the 'Mohole', a convenient and euphonious abbreviation of 'the hole to reach the Mohorovičić discontinuity'.

The feasibility study was concerned with four main topics:

1. How can a ship be kept within a hundred metres of a given point in the open ocean?
2. How can 10 km of drill stem be handled at sea and used to drill a hole to a depth below the sea surface greater than that ever achieved on land?
3. How can the drill be re-inserted in the hole after withdrawing it to change a bit?
4. Where should the hole be drilled?

In the summer of 1960 Bascom issued a report in which all these questions were discussed and no insuperable difficulties found.

DRILLING IN THE DEEP SEA

The next stage was to gain practical experience, to define the difficulties more clearly, and to demonstrate the feasibility of the project by

actually drilling a short distance into the floor of the deep ocean. This was done in April 1961. Most drilling in shallow water has been done from derricks supported on platforms resting on the sea floor. This is not a method that can be used in the deep ocean, most of which has a depth between 3½ and 6 km. In such deep water it is necessary to operate from a floating platform. Drilling rigs have been mounted on various kinds of ships and barges and have been used successfully at sea, particularly off the coast of California. For the preliminary trials for the Mohole the AMSOC committee selected the barge *CUSS I*. The name of this vessel is formed from the initial letters of the names of her owners, the Continental, Union, Shell and Superior Oil Companies.

CUSS I is a converted naval freight barge; she is 79 m long and 15 m wide, and draws 4·6 m when loaded. She carries a derrick 30 m high which drills through a hole in the centre of the ship (see figures 7 and 8). The trial hole was drilled between Guadalupe Island and the west coast of Mexico at a point about 74 km from the island and 220 km from shore (117°30′ W, 29°0′ N); the depth of the water was 3760 m.

When drilling at the test site, *CUSS I* had some 3700 m of drill pipe between her and the ocean floor. If the ship moves, this pipe is bent, and may fracture if the movement is too great. The amount of movement needed to break the pipe is not known, but is believed, from calculation and from experience in shallow water, to be about 200 m. It is therefore desirable to keep the ship within say 60 m of a point directly above the hole being drilled. It is impracticable, or at any rate very expensive, to anchor a large ship with this accuracy in deep water; the varying pulls of wind and current would cause the anchor cables to tighten on one side and to sag on the other, and this would allow the ship to move through some hundreds of metres in an unpredictable manner. It was therefore decided to hold the barge in position, without moorings, by manœuvring her with the aid of four screws, one at each corner. Each of these screws could be rotated about a vertical axis, so that it could exert its thrust in any direction. The combination of the four, therefore, enabled the ship to be moved forward, backwards, or sideways and to be rotated. The speed and alignment of the screws were controlled by a single handle. This was connected to the engines and screws in such a way that the ship moved in the direction in which the handle was pushed. An outer ring round the handle controlled the direction of the ship's head.

These arrangements gave a very convenient and sensitive control of the ship's position and orientation. The marks relative to which the ship was held stationary were a ring of floats attached to totally submerged moored buoys 50 m below the surface (figure 13). A suitable buoy can be moored on a 6 mm steel wire, which it will hold very close to the vertical. The surface float (figure 8) experiences only small forces from wind and current and can be kept within 30 m of the point over the anchor weight. The position of the ship in a ring of buoys is determined by radar and by an acoustic method. Each float carries a radar reflector, which enables its position relative to the ship to be displayed on a 'plan position indicator' visible to the steersman. He sees each float as a bright spot of light on the radar screen and endeavours to keep these spots in marked positions. A similar display is produced by an acoustic system using transponders attached to the buoys. There are thus two entirely separate methods of positioning the ship, both having more than the required accuracy. The whole system worked very well and the ship was kept within 60 m of the required position for many days, in spite of winds of 50 km/h and 4 m waves. This aspect of the problem appears to have found an entirely satisfactory solution, though it is probable that some form of cycloidal propeller will replace the four outboard motors in the larger ship that will be needed to drill the Mohole. The steersman will probably be replaced by an automatic device, in view of the many months of operation that will be needed.

The trial drilling was done with diamond bits with an axial hole. For most of the drilling these were attached to an orthodox drilling rig; the bit was at the lower end of a string of drill pipes 14 cm in diameter screwed together from 9 m sections. 18 m lengths were stowed in racks and raised as required by a pulley block. This block was restrained by guides, which can be seen in figure 9; these are not fitted in a rig for drilling on land, but are necessary at sea to prevent the block from swinging from side to side as the barge rolls.

As the barge rolls and pitches the drill pipe is bent. Calculation shows that the maximum strain occurs near the surface and that, if no precautions were taken, the maximum allowable roll or pitch would be ½°. This will be exceeded in quite moderate weather. The difficulty was overcome by providing a conical guide round the uppermost 16 m of the drill pipe (figures 6 and 14). The guide causes the bend near the upper end of the pipe to be spread over a greater length than if the

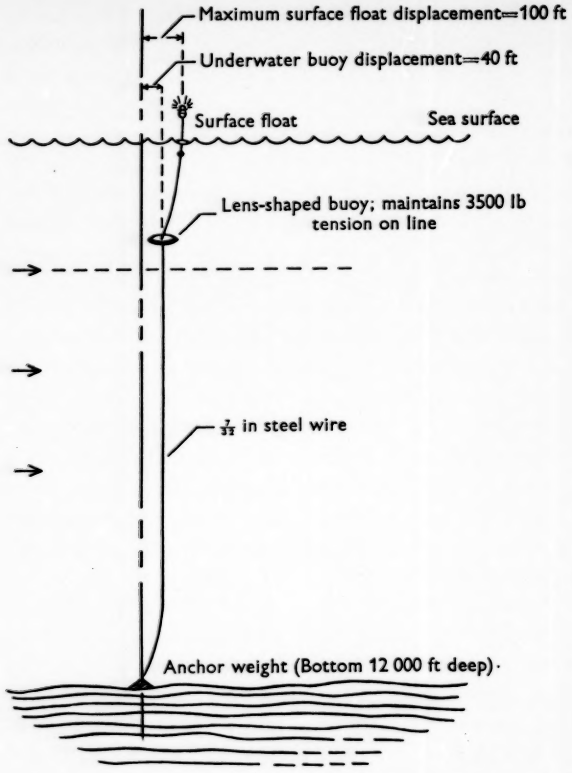

Maximum surface float displacement=100 ft
Underwater buoy displacement=40 ft
Surface float
Sea surface
Lens-shaped buoy; maintains 3500 lb tension on line
7/32 in steel wire
Anchor weight (Bottom 12 000 ft deep)

FIGURE 13 – *A marker buoy. (This figure and figure 14 by courtesy of Weidenfeld and Nicolson Ltd, publishers in Great Britain of 'A hole in the bottom of the sea', by W. Bascom* [1].)

guide had not been present, as is shown in figure 14. With the guide a roll of 6½° can be tolerated.

Chips of rock were removed from the hole by pumping sea water down the drill pipe and letting it flow through the bit, up the hole again, and out at the sea floor. The fitting for the top of the pipe can be seen in figure 10. The bit ground away a hollow cylinder of rock, leaving a central core in a tube on the axis. This tube and the enclosed core of rock could be recovered by pulling them up the drill pipe by a wire (the sandline). The tube could be replaced by carrying it down with the stream of sea water pumped into the drill pipe. It was thus possible to take many lengths of core without withdrawing the drill pipe and bit. When the bit became blunt or some accident prevented further progress, the hole was abandoned and a new hole started. In drilling a deeper hole it would be necessary to be able to re-enter a hole with a new bit.

Some tests were also made with a turbo drill. In this the bit was rotated by a turbine at the bottom of the hole, driven by the stream of sea water pumped down the drill pipe. This method has the advantage that it is not necessary to rotate the drill pipe, and the very large frictional forces between the pipe and the hole are thus avoided. It is also preferable to orthodox rotary drilling in that the diamond bit can be rotated faster, which leads to quicker and more economical drilling.

THE RESULTS OF THE TRIAL HOLES

Three trial holes were drilled close together at the site near Guadalupe Island. The primary purpose of making these holes was to develop methods of drilling in deep water, to find where the problems lay, and to establish the practicability of going deeper. A considerable amount of scientific information was, however, obtained from these holes. The sediments were found to be 180 m in thickness and to be underlain by basalt. The thickness agrees with measurements by explosion seismology, and the finding of basalt is no surprise, since it is by far the commonest igneous rock in the oceans and forms the great volcanic pile of Guadalupe Island only 74 km away. The surprising thing is that the whole 180 m of sediments appears to be of Miocene age, though at the time of writing only a preliminary examination of the core has been made. It is virtually certain that the sea covered this site long before the Miocene; in fact, there is no reason to suppose that it has not been an oceanic area for all past time. It seems highly improbable that the Miocene is the only period in which sediment was deposited, and we must therefore conclude that other sediments are concealed beneath the basalts of the Pacific. There is nothing unreasonable or surprising about this, but there had been no evidence for it previously; it does ease the problem posed

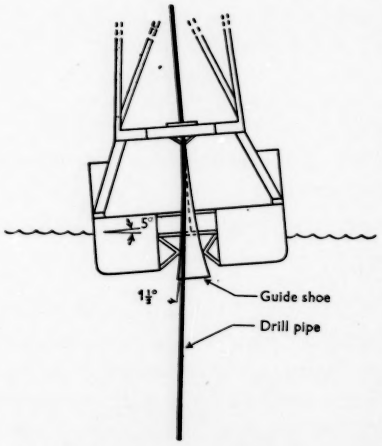


FIGURE 14 – *The effect of rolling with drill pipe in conical guide.*

by the small thicknesses of unconsolidated sediment found all over the world by the seismic method. We can now expect that much of the material between the Moho and the ocean floor will be found to be composed of sediments interlayered with basalt. The total length of time represented by the Miocene period is usually estimated at 14 million years; the sediments penetrated represent less than the full span of the period. If they are assumed to have been deposited in 5 million years, the rate of sedimentation is 4 cm/1000 yr, which is several times the average rate determined from short cores. Why sediment should have accumulated rapidly at this point during the Miocene and very slowly, or not at all, in the recent past, is not known.

The rock beneath the sediment is a fresh-looking crystalline basalt with a millimetre or two of glass at its upper surface. So far as the writer could see, there was no baked sediment sticking to it; it is therefore to be supposed that it flowed out on the ocean floor and was not intruded as a sill into pre-existing sediment. The basalt was penetrated for 15 m and a large amount of core collected. The results of the examination of this material will be awaited with interest. It will presumably be possible to determine its age, as well as its chemical composition and physical properties.

Temperature measurements in one of the holes gave a heat flow about twice the average for the oceans, this high value being in agreement with measurements made in the vicinity with short probes. These measurements give a most welcome confirmation of the results from the probes, since, in spite of the care with which possible errors in the probes have been investigated, it was not possible to feel entirely certain that measurements over a range of depth of a few metres gave really significant temperature gradients.

PROBLEMS FOR THE FUTURE

After the successful completion of the trial holes we may expect future work to divide into two streams. First, with the existing technique or slight extensions of it, and with ships the size of *CUSS I*, holes up to 300 m can certainly be drilled into the material beneath the floor of most parts of the ocean. This opens the possibility of studying in detail the sediments and lavas to that depth and of following the history of the oceans for a very considerable period in the same way that it has been followed on land. This should lead in a few years to the solution of some of the major problems of geology. It will be possible to establish beyond doubt whether the land masses that have been supposed to exist off many continental coasts have foundered in the present oceans; it is the expectation of the writer that such submerged continents will not be found. It should also be possible to find the age of each ocean basin and, from changes in the sediments, to study past changes in the distribution of land and sea, and perhaps changes of climate and ocean currents. The cost of such a programme would not be high. *CUSS I* cost about three million dollars, and it would probably be possible to operate a similar ship for something under one million dollars per year. It would seem well worth building at least one such ship instead of one or more of the proposed new conventional oceanographic ships.

The other line of development will be the attempt to reach the Moho. Since so deep a hole cannot be drilled with a single bit, it will be necessary to withdraw the drill pipe and bit, and re-enter the same hole with a new bit. This is an operation that has not been tried in deep water. There are at least two possible methods. A pair of wires might run from the ship to the bottom carrying a horizontal bar between them. If the lower end of the drill pipe were attached to the centre of this bar, it could slide down the wires and direct the bit into the hole. This method has been used in shallow water. If it were used in deep water, there is some danger of the wires and the drill pipe becoming entangled. Also it would not provide a return path to the ship for fluid pumped down the hole, and the drilling fluid would therefore have to be sea water, as in the trial holes. In drilling on land it is usual to use a thixotropic mud as a drilling fluid, and the impossibility of doing so on the ocean bed is a grave disadvantage of the scheme. Mud serves several purposes in drilling. It becomes plastered on the side of the hole and prevents material falling from it; if the mud is of high density it prevents water or oil in the rocks from being forced violently upwards and out of the hole; and it also prevents the rock chippings from falling back on the bit if circulation of the mud is stopped.

It seems imprudent to adopt a method that prohibits the circulation of mud, and there would be a great advantage in installing a pipe between the sea floor and the ship, down which the drill pipe could run, and up which the mud could flow back. It would be necessary to support this pipe with submerged floats so that it could be abandoned by the ship in bad weather or other emergency and later reconnected. If such a pipe is to

be transported and installed, a ship considerably larger than *CUSS I* will be required. During the preliminary trials a prototype of the bottom support for such a pipe was installed in relatively shallow water off San Diego, but so far no tests have been made in deep water. There are considerable difficulties in handling, assembling, and maintaining so cumbrous an object as a pipe perhaps 30 cm in diameter and 4000 m long. It seems likely that further trials will be necessary before drilling of the Moho is attempted. A turbo drill with a diamond bit appears to have great advantages for a very deep hole, and will probably be used.

No site for the first Mohole has yet been selected. It should be in a 'typical' ocean floor, whatever that may mean. The Moho should be as near the surface of the ocean as possible and the heat flow should not be too high, though it is probably not necessary to seek a place where it is specially low. It is not clear whether, for a given depth of Moho below the surface of the ocean, very deep water is an advantage or a disadvantage. The weather at the site should be good for a large proportion of days in all seasons, and the site should not be too far from sources of supply of equipment or from a congenial port in which the crew can spend their periods of leave.

Considerations of weather exclude the Atlantic and the Pacific north of about 40° and also the areas of the trade winds and the hurricane belt. The heat flow off the southern Californian and Mexican Pacific coasts is higher than is desirable, but the good weather and the logistic advantages make a site somewhere in this area a likely choice.

The cost is hard to estimate, but it is probable that, before the first hole is finished, about fifteen million dollars will have been spent; the cost is unlikely to exceed twice that sum.

THE RESULTS TO BE EXPECTED

One or more holes penetrating below the Moho would resolve many uncertainties. They would show whether the Moho is a boundary between two types of rock or is a place where a change in crystal structure takes place. The rate of heat generation by radioactivity in the outer part of the mantle could be measured, which would enable us to know the limits of the wide range of temperature distributions that might, for all we at present know, occur within the earth. In particular, if the material below the Moho were found to resemble the stony meteorites in chemical composition and radioactive content, a fairly detailed account could be given of the earth's thermal history.

These things are fairly specific; perhaps the main reason for drilling the Mohole is a more general one. We have no firmly based knowledge of the history of the oceans. Within the last year papers have appeared maintaining (*a*) that the Atlantic ocean widens by continually splitting down the middle, (*b*) that all the oceans widen by a general expansion of the earth, and (*c*) that the material from the ocean floor is continually swept beneath the continents by stresses due to convection currents in the mantle. These are the views of the *avant-garde* of geology, but there are still many who believe in the permanence of the oceans, in their production by the foundering of continents, and in their conversion into continents.

These problems and most of these views have been with us for many years. The only hope of discriminating between the innumerable possibilities is to drill in the ocean floor and find what is in fact there and what is really happening. The Mohole is one of the greatest of scientific endeavours. Its achievement should lead to an outstanding advance in our knowledge of the interior of the earth.

ACKNOWLEDGMENTS

AMSOC and Dr Bascom allowed me to visit the trial drilling off Guadalupe Island last April. The ordered chaos that I saw there deeply impressed me, not only as a scientific project, but as a co-operative venture between very different kinds of people, sailors, drillers, oceanographers, and engineers among others. Without this visit I could not have written this article. I am also indebted to Dr Bascom's recent book [1] for much information.

REFERENCES

[1] BASCOM, W. 'A hole in the bottom of the sea'. Doubleday, New York. 1961.
[2] FERMOR, L. L. *Spec. Publ. Indian Ass. Cult. Sci.*, No. 6, 1938.
[3] HEEZEN, B. C., THARP, M., and EWING, M. *Spec. Pap. geol. Soc. Amer.*, No. 65, 1959.
[4] LOVERING, J. F. *Trans. Amer. geophys. Un.*, **39**, 947, 1958.
[5] MOHOROVIČIĆ, A. *J. met. Obs. Zagreb*, **9**, Pt. 4, Sect. 1, 3, 1909.
[6] WOOLLARD, G. P. *J. geophys. Res.*, **64**, 1521, 1959.

Light quanta and vision

M. H. PIRENNE

The sensitivity of the eye is so great that it can be stimulated by a very few quanta of light. At luminances near the threshold of vision the quantum properties of light have important consequences, and they are also important in the general theory of visual processes. This article discusses some of the relevant phenomena and their theoretical interpretation. It considers also the retinal activity that may occur in total darkness and draws attention to the role of the nervous responses that occur when light is cut off from the retina.

Considered as a detector of electromagnetic radiation, the human retina has a very limited wavelength range of sensitivity, namely the visible range of the spectrum. Within this range, the retinal sensitivity[1] depends very much upon wavelength (figure 2). The sensitivity is highest when a wavelength in the neighbourhood of 500 mμ is used under such conditions that vision is mediated by the rod system found in the periphery of the retina. The numbers of quanta acting directly upon the retinal rod receptors and stimulating them are then often so small that the quantum properties of radiation have to be explicitly taken into account in the theory of visual processes, particularly when considering the purely physical fluctuations that occur in the light stimulus itself.

All the available evidence indicates that the decrease in the sensitivity of the rods for wavelengths shorter or longer than about 500 mμ is due to the fact that the photosensitive substance contained in them (rhodopsin or visual purple) absorbs proportionately less light at these wavelengths. Accordingly, the numbers of quanta actually caught by, and able to act upon, the rods remain very small when monochromatic light of wavelengths other than 500 mμ are used, and when light with a broad spectral distribution, such as white light, is used. Thus, when investigating the smallest quantity of light that will just suffice for seeing with rod vision, it is found that about a million times more quanta incident on the retina are required for light of 720 mμ than for light of 500 mμ; yet it is most probable that the number of quanta absorbed by the rods is essentially the same in both cases.

Experiments show that the quantity of light energy just sufficient for a visual response is smallest in the case of a brief flash of light of about 507 mμ observed by peripheral vision, the eye having previously been fully adapted to the dark. The energy content of one such flash is about 100 quanta; this is measured at the cornea and represents the mean amount of energy entering the eye. It is to be noted that the quantum is used here merely as a convenient unit of energy, since the quantum properties of light come into play only when light is absorbed, quanta being impossible to localize in a flux of radiant energy. Mean energy values will sometimes be expressed as fractional numbers of quanta even though, of course, the number of quanta absorbed is always an integral number. As will be shown later, the mean number of quanta absorbed by the rods in these experiments is about ten; it may possibly be less than ten, but always exceeds one. These quanta are distributed among a certain number of rods (a few hundred rods for a test field subtending an angle of 10′ of arc at the eye (figure 3)). It is impossible, if only because of the diffraction that occurs at the pupil, to confine the stimulus to one rod only. The minuteness of the quantities of energy involved in these experiments is illustrated by the fact that the mechanical energy required to lift a pea one inch would, if converted into light, be sufficient to show a faint light signal to all the men who have ever lived.

The ability of the eye to respond to the absorption of very small numbers of quanta by the retinal receptors is of practical importance in everyday life. For instance, if we wake up in the middle of a very dark night and look at the window of our bedroom, we can distinguish features of the window against the lighter patch formed by the sky. If, after a minute or two, we again close

[1] Spectral sensitivity, as used here is the reciprocal of the threshold quantity of light, the latter being defined as $W_\lambda \tau_\lambda / h\nu$, where W_λ is the threshold energy measurement at the cornea, τ_λ is the transmission factor of the eye media, and $h\nu$ is the quantum energy. This is the sensitivity at the retinal level on a quantum basis. The sensitivity measured at the cornea of the eye on an energy basis is $1/W_\lambda$; for the rods it is maximum at 507 mμ [31, 32] instead of near 500 mμ.

our eyes, our retinas during the whole operation will not have received enough light to enable each of the receptor cells involved to absorb even one single quantum of light. In a more general way, we may see quite well enough on dark nights to find our way about, even though at the prevailing level of illumination the majority of our retinal receptor cells remain at any instant wholly unstimulated, the sensitive substance they contain absorbing no light whatsoever. Only at illumination levels corresponding to full moonlight is the intensity enough to make each rod absorb several quanta per second.

THE ROD AND THE CONE SYSTEMS IN HUMAN VISION

The retina contains two types of receptor cells, the rods and the cones (figures 3 and 4). It is the rod system which is mainly active in night vision and in all conditions when the eye is responding to low illuminations to which it has become fully adapted. After stimulation by bright daylight, it takes at least half an hour for the eye to become fully dark-adapted.

Some of the properties of the rods have already been described. The photosensitive pigment, rhodopsin, is contained in the terminal segments of the rod cells (*b* in figure 4). It is pink in colour, because it absorbs most strongly in the blue-green region and transmits the red end of the spectrum almost freely; rod sensitivity shown in curve 1 of figure 2 parallels light absorption by rhodopsin in the range where the latter can be measured [33]. In the case of dark-adapted retinas of man or of such animals as the frog, in the fresh state, this pink colour can readily be seen under the microscope under weak illumination. Under strong illumination, the colour fades rapidly as the rhodopsin is bleached by the light. It is possible to prepare aqueous extracts of rhodopsin from retinas, and its properties have been extensively studied *in vitro*.

Only quanta absorbed by the photosensitive substance can affect it; needless to say, fractions of quanta cannot be absorbed. The latter fundamental fact often simplifies the discussion of the action of light stimuli upon the retina. For instance, when the amount of energy absorbed corresponds to a number of quanta small compared with the number of receptors involved, it follows that some of the receptors will have absorbed one or several quanta, whereas the others will have absorbed no light at all. It is not quite certain that all the quanta so absorbed are 'acting quanta' in the sense that they lead to nervous excitation of the receptor cells; but in any case, the number of acting quanta can only be equal to, or smaller than, the number of absorbed quanta. It is generally accepted that what matters in rod stimulation is the number of light quanta acting upon the rods, not the energy of the individual quanta. We shall later discuss experiments made with light of about 507 mμ, but, in terms of the number of quanta acting upon the rods, the results must also be valid for other lights, for instance the light from the night sky.

The cone system of the eye is active in normal daylight, giving us then sharp and coloured vision. On account of the fixation reflex, we use direct or foveal vision to look at any point of an object which attracts our attention. The very centre of the *fovea centralis* of the retina contains cones only (figure 1). It becomes blind at low illumination levels at which the rod system, present only in the periphery of the retina, is still able to function. This is why we must—in spite of the fixation reflex—look 'out of the corner of our eye' in order to see best at night. The astronomer Arago said epigrammatically that in order to see the faintest luminous objects one must not look at them. This lower sensitivity of the foveal cone system is due in part to a smaller capacity for summating light stimuli in space and in time; this in turn gives the fovea greater accuracy of vision than the periphery. Even when small and brief test-stimuli are used, however, the smallest number of quanta required for seeing, measured at the pupil, is found to be markedly greater for the fovea than for the periphery. The foveal cone system, considered as a whole, has its maximum sensitivity near 555 mμ. It contains several cone-mechanisms having different spectral sensitivities [3]. The individual cones themselves must vary in their properties, and most probably they contain several different sensitive pigments [4]. Much less is known about the cone pigments, particularly with regard to the amount of light they absorb *in situ*, than about rhodopsin [5]. Consequently, the study of quantum effects is more difficult and less advanced for the cone than for the rod system, and in this article we shall concentrate our attention on the latter.

Whereas there are only cones in the centre of the fovea, all other retinal regions contain both cones and rods; in the periphery, however, the rods greatly outnumber the cones. Except in the case of long wavelengths, at the red end of the spectrum, vision with the dark-adapted periphery

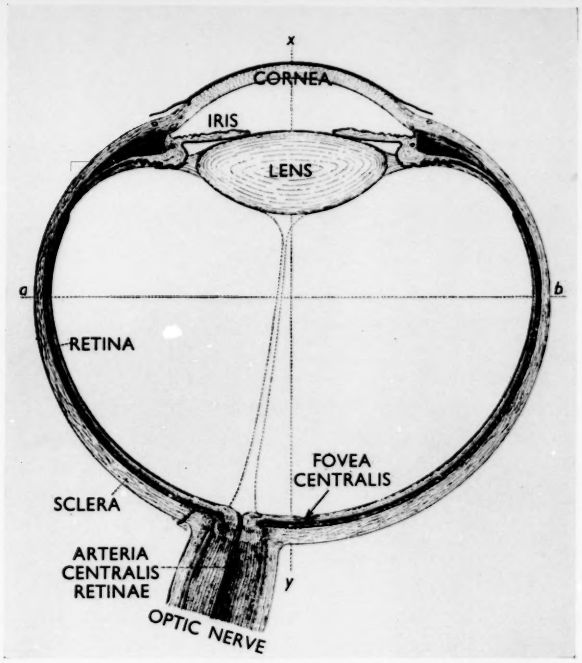


FIGURE 1 – *Section of an adult human eye. The* fovea centralis *is a depression of the retina, only a fraction of a millimetre in diameter, situated near the optic axis* xy *of the eye. Its centre contains cones only. The rest of the retina contains both rods and cones.* (×3) (*From E. A. Schäfer, 'Quain's Elements of Anatomy', Vol. III, Pt. III. London,* 1894.)

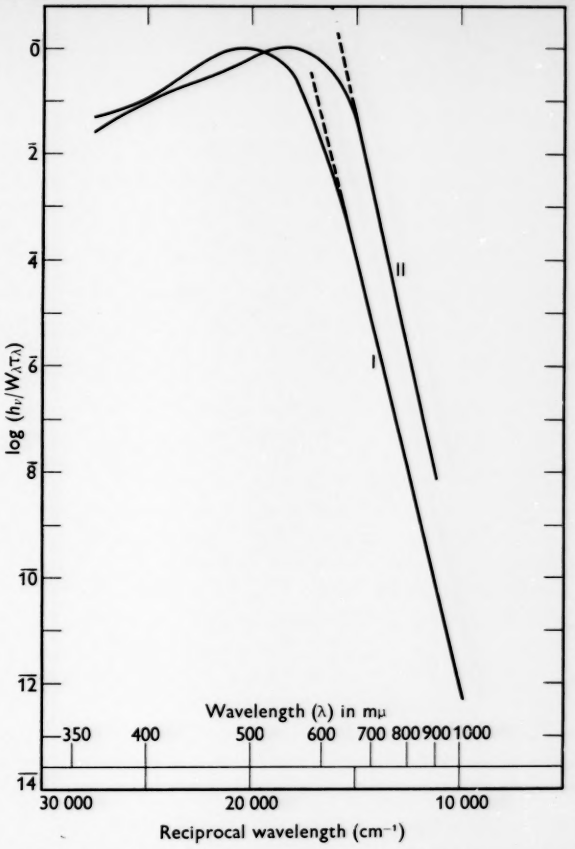


FIGURE 2 – *Relative spectral sensitivity of the human rod and cone systems. Curve I, parafoveal rod vision. Curve II, foveal cone vision. In the figure the maximum values of the sensitivity curves for the rods and the cones have been arbitrarily set equal; their actual separation on the vertical scale varies with the experimental conditions. (After W. S. Stiles,* [1].)

at the lowest intensities is mediated by the rod system alone. This has been demonstrated in experiments using the 'directional sensitivity of the retina' (Stiles-Crawford effect [5–7]) which is almost exclusively a property of the cones, not of the rods (figure 5). Roughly speaking, the cones are much less effectively stimulated by rays falling on to the retina obliquely instead of normally, whereas rod stimulation is, as a rule, almost independent of the direction of the incident light, within the range of directions that can be obtained using the natural pupil. Deep-red light often stimulates the dark-adapted peripheral cones as well as the rods, but by the use of a narrow beam of red light of suitable intensity, entering through the edge of the pupil, stimulation can be restricted to the rods. Suppose that in figure 5 the 'red' rays X and Y, of equal intensity, fall on the dark-adapted retina at the same spot Q, containing both rods and cones. X and Y stimulate the rods almost equally. At fairly high intensities, the ray Y, because it falls normally on the retina, stimulates the cones more than X, which falls obliquely. Now the intensity can be reduced so that, while Y still stimulates both cones and rods, X stimulates rods only. If a pencil of red light is

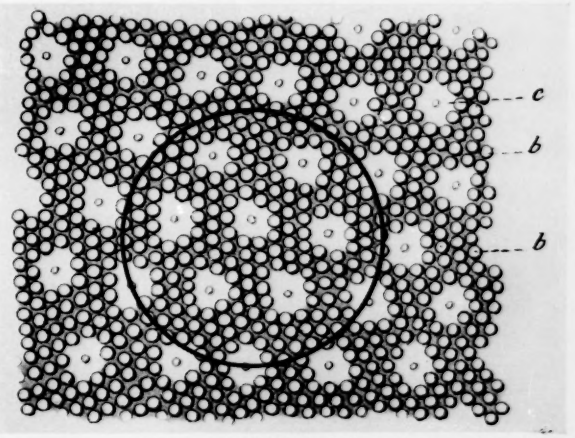


FIGURE 3 – *The mosaic of rods* (b) *and cones* (c) *in the periphery of the human retina* [2]. (×750) (*From M. Schultze,* 1866.) *The large circle represents the geometrical image,* 48·5 *microns in diameter, of a test field subtending a visual angle of* 10′ *of arc.*

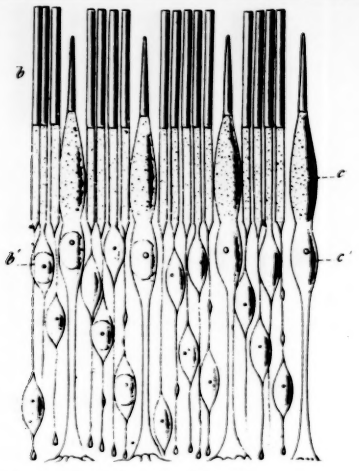


FIGURE 4 – *Section through the layer of rods* (b) *and cones* (c) *in the periphery of the human retina. The light from the pupil passes through the retina and finally reaches the rhodopsin contained in the terminal segments of the rods* b, *drawn darker in the figure.* (×500) (*From M. Schultze* [2].)

made to traverse the diameter of the expanded eye pupil, the 'red' source is first seen as almost colourless, whitish or greyish, then as red, and finally again as colourless. That is, for peripheral pupil-entry, long-wave light gives the same almost colourless sensation as given by light from the other parts of the spectrum for both peripheral and central entry. Again, the relative spectral sensitivity is found to be the same for central and peripheral pupil-entry throughout the whole spectrum except for the red region [8]. The red colour often seen with the dark-adapted periphery in the case of long-wave light stimuli, of sufficient intensity, must therefore be ascribed to the action of the cones. When the rod system is stimulated in isolation it is incapable of discriminating between different wavelengths, a finding which, among other things, gives support to the view that, within the visible range, all acting quanta cause nervous stimulation of the same kind in the rod cells, independently of the energy of each quantum. All this is related to the well-known observation that on a dark night it is impossible to distinguish the colours of objects. Red objects then look neither red nor greyish, but black, because they do not reflect enough light to stimulate either the cones or the rods.

SENSITIVITY OF ROD VISION

One of the simplest illustrations of the sensitivity of the rod system is provided by the absolute threshold of the eye for a large illuminated test-

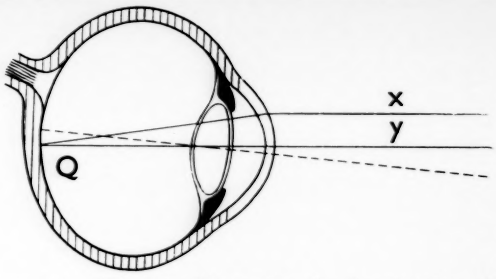


FIGURE 5 – *Diagram illustrating the directional sensitivity of the retina (Stiles-Crawford effect).* (*From W. S. Stiles,* [7].)

field emitting a band of wavelengths near the 507 mμ region and presented in long exposures. The conditions in one set of experiments were: field diameter subtending an angle of 47° at the eye; exposures each lasting about 15 seconds; and binocular vision with free eye-movements [9].

The absolute threshold is, broadly speaking, the lowest intensity at which the test-field can just be seen. This value must be defined on a statistical basis, however, because, for any given intensity-setting of the apparatus near the limit of visibility, the field is sometimes seen, sometimes not seen. There is no sharply defined setting below which the light is never seen and above which it is always seen. There is instead an intensity range where seeing is uncertain and unpredictable. Such a range of uncertain seeing is encountered in all experiments on visual thresholds. It must be explained by a combination of biological variation within the organism and purely physical fluctuations in the light stimulus itself. In the case of a small brief flash presented in the periphery it has been possible to show that the quantum fluctuations, which are then relatively very large, because the number of quanta acting upon the retina is very small, are a major cause of the uncertainty of seeing [10–14].

Absolute threshold is, by definition taken as the intensity at which the presence of the large illuminated field can be detected by the eye in 50 per cent of the trials, each trial lasting 15 seconds. There are variations from subject to subject in the threshold-value so defined, and even from day to day in the same subject. Similar variations are observed in the case of the statistically defined threshold-values corresponding to other experimental conditions, for instance, a small brief flash.

The average threshold-value so obtained for young normal subjects leads to the conclusion that, at the limit of visibility for our large field, each rod absorbs on the average only one quantum

every hundred minutes. The energy of the quantum for $\lambda = 507$ mμ is $3{\cdot}92 \times 10^{-12}$ erg.

This conclusion rests, firstly, on a straightforward measurement, that of the mean value of the absolute energy of the light flux entering the eye through the pupil at the threshold value. This is made by determining the energy flux from the source at a higher intensity and cutting it down in the threshold measurements by the use of filters of known transmission properties. The second main basis of the above conclusion is an estimate of the fraction k of the light of 507 mμ wavelength that, having entered the eye, eventually becomes absorbed in the rhodopsin contained in the living human rod-cells. The value of k is derived (*a*) from a determination of the light losses occuring in the optical media of the human eye, from the cornea to the retina, and (*b*) from a determination of the fraction of light absorbed by the rhodopsin in the living rods. Both determinations suffer from some uncertainties, and, of course, the values probably vary to some extent from eye to eye. The fraction of the light of 507 mμ wavelength which reaches the retina through the eye media is 0·50 according to E. Ludvigh and E. F. McCarthy [15]. The fraction of this light which is then absorbed by the rhodopsin in the living rods has been estimated by W. A. H. Rushton [16] to be 0·20, on the basis of his delicate method of reflection densitometry. This is essentially a method by which the intensity of the light reflected from the back of the retina is measured before and after completely bleaching the rhodopsin *in situ* with a strong light. The difference therefore corresponds to the light absorption in a double passage through the layer of rhodopsin contained in the dark-adapted rods. On the above assumptions $k = 0{\cdot}50 \times 0{\cdot}20 = 0{\cdot}10$.

The exact value of the fraction of light absorbed by the rods is of relatively minor importance in the present context. Even on the extreme, and impossible, assumption that 100 per cent of the light entering the eye is absorbed by the visual substance contained in the retinal rods, the absolute threshold for the large field considered would still have the remarkably low mean value of one quantum absorbed per rod every ten minutes.

The value $k = 0{\cdot}10$, which should be considered as a reasonable estimate rather than as an accurate measurement, will be taken here as a basis for further discussion. It leads to a mean of $1{\cdot}73 \times 10^{-4}$ quantum absorbed per rod per second at the absolute threshold for our large field and long exposures. On the average, therefore, one quantum is absorbed every second per 5800 rods. The

FIGURE 6 – *The Landolt ring, a test object for the study of visual acuity. The width of the gap in the ring and the thickness of the black ring are equal to one-fifth of the total diameter of the ring.*

retinal image of a circular test-field subtending 47° at the eye has an area of about 1·5 cm². When this image falls on the retinal periphery, it covers about 750 000 cones and 20 million rods. This large number of rods absorbs an average of only 3500 quanta per second under the conditions just described.

ROD VISION AT HIGHER LUMINANCES

At the absolute threshold, the luminous field, in those trials in which it is seen at all, has an extremely vague, diffuse appearance. But if the

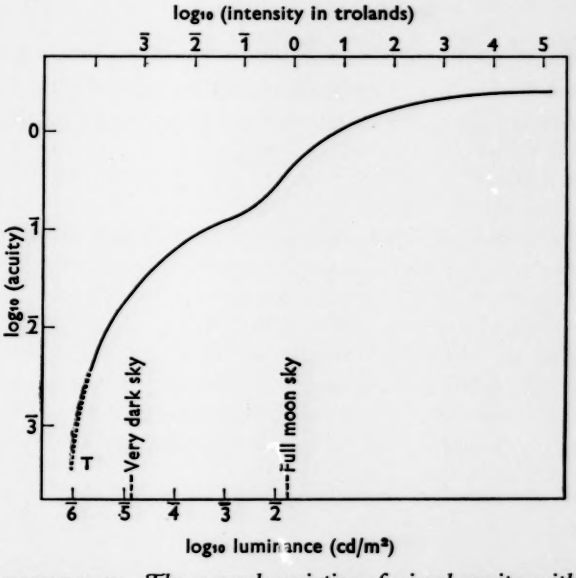


FIGURE 7 – *The general variation of visual acuity with light intensity for a black Landolt ring. Visual acuity is defined as the reciprocal of the angle, in minutes of arc, subtended at the eye by the gap in the ring. One troland is the retinal illumination produced by a surface having a luminance of* 1 *candela per square metre when the pupil area is* 1 *square millimetre. The external field luminances for the lower intensities are given on the lower abscissa scale. The point* T *indicates the average value of the absolute threshold luminance for an empty area, of* 47° *diameter, for young subjects.* [*See* 9].

light intensity is raised to three times the threshold value, the subject sees the field clearly, and can distinguish the fingers of his hand in front of it. If a black Landolt ring (figure 6), the gap of which subtends an angle of one degree at the eye, is placed in the middle of the field at ten times the absolute threshold intensity, a trained subject can detect the position of the gap in 50 per cent of the trials. Thus form-vision is here already acquiring a measurable degree of efficiency, even though the mean light-absorption per rod has only been raised to about 0·002 quantum per second—that is, one quantum every ten minutes or so. This level of intensity is a little lower than that which corresponds, in terms of quanta absorbed by the rods, to the darkest night skies [9] (figure 7).

Because the illumination values quoted above are so low in terms of acting quanta, the quantum nature of light retains its importance at higher intensity levels, where vision becomes relatively efficient. The luminance of the sky at full moon is roughly 20 000 times that corresponding to the absolute threshold of the eye. Here the position of the gap can be detected in a black Landolt ring of much smaller angular dimensions, namely when the gap subtends 3 minutes of arc at the eye. This is a fairly good degree of resolution, since the maximum resolution that the eye can achieve in bright light for the same type of object is only seven times greater, corresponding to a gap of about 0·4 minute of arc. (This maximum resolution is achieved with the central foveal cones, and corresponds roughly to the angular separation of these cones [34].) In the case of the moonlit sky, each rod on the average absorbs about 4 quanta per second. In the primary photochemical reaction, each of these quanta acts upon a single molecule, or rather upon a single chromophoric group, of rhodopsin. The number of such chromophoric groups contained in one human rod is of the order of 10^8. The quantum yield of rhodopsin bleaching is not higher than unity: that is, the ratio of the number of rhodopsin molecules that become bleached as a result of a quantum absorption to the number of molecules that have absorbed a quantum, may be unity, or possibly one-half, but not greater than unity. There is no chain reaction whereby one quantum absorbed might lead to the bleaching of several molecules; this conclusion is mainly derived from measurements of quantum yield made on solutions of rhodopsin and on rhodopsin, *in situ*, in animal retinas. Thus, as E. Baumgardt has emphasized [17], the rate of bleaching of rhodopsin in the rods remains very small over the whole range of intensity levels corresponding to night vision. At the luminance of the moonlit sky, it reaches only 4×10^{-8} per second, or 0·015 per cent per hour. Clearly the action of light upon the rods is not an ordinary photochemical reaction involving large numbers of molecules; the absolute number of rhodopsin molecules acted upon by light in each rod in night vision is small. An amplification mechanism, the nature of which is as yet unknown, must be at work in the rod receptor cell. At moonlight levels, parafoveal cones probably co-operate with rods. At higher luminances, the foveal centre becomes the most accurate part of the retina, and the cones take over.

When rod vision is effectively used, it always relies on small numbers of quanta per rod, and the total number of rhodopsin molecules contained in each rod is barely decreased by bleaching, even at the higher level of the illumination range that normally corresponds to rod vision. Towards the lower end of this range, as we have seen, the fraction of the rods actually stimulated by light per unit time is very small. The appropriate time unit is probably not less than 0·1 second; this is the approximate value of the physiological action time derived from experiments on temporal summation in absolute threshold experiments. For exposures shorter than, or equal to, this physiological action time, the mean total amount of light delivered to the eye is the determining factor. For longer exposures, the total amount of light required for threshold response increases. Assuming that the effect of one quantum on one rod cell also lasts 0·1 second, it may be said that in the case of extended fields—presented in long exposures, say of the order of one minute, at low luminances—the rods receiving the 'retinal image' of the field are for most of the time completely unstimulated. If, on the average, one rod absorbs one quantum every ten minutes or so, corresponding to a very dark night sky, it will be in a state of excitation for only about 0·1 second in every ten minutes. Certain experiments, however, suggest that light absorption by the retina may, even at intensities near the absolute threshold, lead to certain types of physiological reactions lasting as long as 5–10 seconds [18]. Rod stimulation by one quantum might last as long as this, but even so it would only occur during a fraction of the time at low intensities. As will be discussed later, experiments made with long exposures and with the eye free to move in the ordinary way probably bring into play complex physiological mechanisms. For this

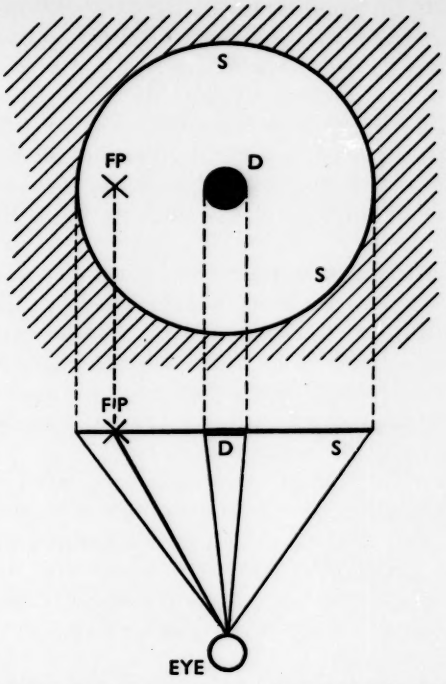


FIGURE 8 – *Arrangement for experiment by the fixation-and-flash method.*

reason, experiments made with brief flashes, shorter than 0·1 second, are often easier to interpret than experiments made with long exposures, but they no longer correspond exactly to the conditions of ordinary vision.

ACUITY OF ROD VISION, AND ITS DEPENDENCE UPON THE LEVEL OF ILLUMINATION

The familiar fact that visual acuity—that is the resolving power of the eye, in terms of form vision, or of the ability to see details—becomes higher when the light intensity increases, for such test objects as the black Landolt ring, has long remained without any adequate explanation. Some might be tempted to dismiss this problem by saying that it is obvious that we could not see well when there is not enough light; on the basis of the quantum theory, there is in fact a great deal of truth in this apparently tautological argument. For at low intensities the smallness of the number of quanta acting upon the receptors must set a natural lower limit to the size of the smallest black object, or detail of an object, that can be seen at a given intensity level. The actual limit does not seem to be very much higher than this limit, set by the properties of light itself.

In some experiments on rod visual acuity, made

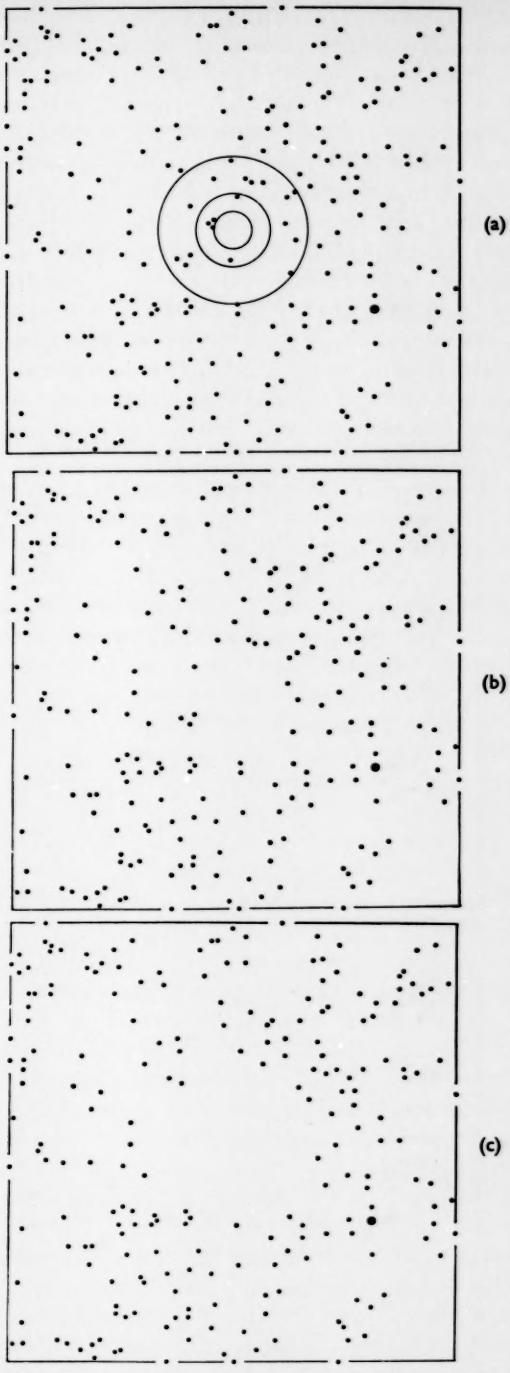


FIGURE 9 – (*a*) *Calculated random distribution of* 250 *points in a region divided into* 10 000 *small equal areas. (The large black spot indicates two points that have fallen into the same area.) The large concentric circles in the middle of the diagram are meant to represent the extent of the retinal images of three black disks of increasing diameters.* (*b*) *and* (*c*) *The same distribution of quanta as would be effected by the presence of the dark images of the two larger disks.*

with a black disk presented in the middle of a wide illuminated field, the task of the subject was merely to detect the presence of the disk, which is one of the simplest objects that can be used to test acuity of vision. We have discussed the more complicated case of the Landolt ring, partly because it is a typical test-object of the kind used for measuring visual acuity in the prescription of spectacles. With the help of a fixation point (figure 8), the luminous field, together with the black test-object, was presented always in flashes of a few milliseconds duration. The subject fixed his gaze on a small red point of light FP that was just comfortably visible in the dark. Since the image of the fixation point FP was always formed in the middle of the fovea, the retina was fixed in space —apart from inaccuracies of fixation, amounting to a few minutes of arc, that are not important in the present experiments. The subject operated a shutter and reported whether or not he had seen the black disk D—which for checking purposes was in some trials removed from the field without the subject's knowledge. These experiments were made with the dark-adapted eye, the eye being allowed to recover its dark-adaptation between flashes. Consequently, the experiment as a whole refers to a constant physiological state of the visual system.

The apparatus is arranged so that the wide field, S, surrounding the black disk, D, is evenly illuminated, yet the quanta actually absorbed by the rods in each flash are randomly distributed and form in the retina a pattern of the kind shown in figure 9. The random retinal pattern of quanta is of course different in each flash, though for convenience the same pattern has been used in all the diagrams of figure 9. It is clear that the dark image of a black disk will not be detected if its diameter is small compared with the mean distance separating neighbouring absorbed quanta. In this case, the random patterns of absorbed quanta formed when the disc is present will in practice be indistinguishable from the random patterns which would be formed when there is no black disk in the middle of the field. For each flash there are, between the rods which have absorbed a quantum (or sometimes more than one quantum), retinal areas completely unstimulated by light. The presence or absence of a dark 'shadow' on these areas cannot be detected in the experiment.

Theoretically, detection of the black disc should become possible when it produces on the retina a shadow larger than the unstimulated regions occuring naturally between the absorbed quanta. The situation may be illustrated by comparison with a dry road on which a few scattered drops of rain have fallen. If a large object, such as a motor car, is then moved away, its previous position will be detectable by the absence of raindrops, but this will not be so if a small object, such as a pebble, is picked up from the surface of the road.

Thus the quantum pattern of retinal excitation sets a natural limit to the accuracy of vision as studied in such experiments. And it must be expected that smaller black objects will become detectable when the flash intensity is higher, because the mean number of quanta absorbed per unit area of the retina is then larger.

Experimental results do agree qualitatively with this prediction, but it is generally accepted that one acting quantum is not sufficient for the eye to detect a light stimulus. The theoretical limit to the accuracy of vision, which has just been discussed, is an upper limit, which we cannot in fact expect to reach.

If one acting quantum were sufficient for vision, the light stimulus would be seen when at least one quantum had been absorbed by, and had acted upon, one rod anywhere in the retina. Statistics show that the probability of this occurring depends solely upon the total mean number of quanta absorbed by the rods and is independent of the retinal area or areas involved and of the distribution of the stimulus in time. Accordingly, if one quantum were enough for seeing, the result of absolute threshold experiments (detection of a light stimulus presented on a zero-intensity background) would be the same as if there were complete 'physiological summation' (see later) by the visual system in space and in time; in fact summation is found to occur only over a limited area and a limited time. The mean total amount of light required for seeing, as defined by the setting of the apparatus and the conditions of the experiment, is not independent of area and time. Again, it is greater when two eyes are used instead of one [19], whereas for 'one-quantum vision' it would be expected to be about the same for one eye as for two eyes, even when the two eyes belong to two different human subjects responding independently of one another [13].

We must conclude that several acting quanta are required for vision, and that these quanta must be absorbed within a limited retinal area. They must also be absorbed within a period of less than about 0·1 second—a condition which is fulfilled

when brief flashes are used, as in the acuity experiments described above.

Spatial summation in the retina can be studied by using as a test object a pair of light spots placed at various distances apart. If these spots are separated by an angular distance roughly 20′ of arc or less in the periphery, their effects are to some extent additive and there is at least partial physiological summation. When the two spots are sufficiently near to one another, the threshold is determined solely by the total amount of light delivered by the two spots; this is complete physiological summation. If the two spots are far apart, there is still a greater chance of seeing at least one of the spots when both are presented together than when they are presented singly. The same increase in the probability of seeing would occur if a single spot were presented to two independent subjects, the spot being considered to be seen when at least one of the subjects has perceived it; in this case there is only 'probability summation' and no physiological summation. At the threshold, the two eyes of a subject respond very nearly on the basis of probability summation [19].

Thus there is a limited range of physiological summation of the effects of light quanta absorbed in the dark-adapted retina. Instead of considering single quanta, or rather, single rods each having been acted upon by at least one quantum, in the pattern of stimulation we must consider functional 'visual units' each containing many rods. For each unit, these rods might simply be all those that are connected to a single optic-nerve fibre (figure 12); it is also possible that the effective convergence occurs at other levels of the nervous system, such as the brain. Each of these units will be stimulated when it has absorbed a certain number of quanta, which is generally accepted to be at least two and probably not more than about ten.

Assuming, for the sake of simplicity, that we are working with a retinal region having uniform properties over the whole of its area—an assumption which is never strictly fulfilled for retinal areas of the extent considered here—a theoretical interpretation of these experiments may be attempted by considering that the dots of the diagrams of figure 9 correspond, not to single absorbed quanta, but to independent visual units, each of which has absorbed enough quanta to become excited.

The relevant point now is how many visual units have received multiple hits of sufficient magnitude to be stimulated. As long as there are, in the retinal area receiving the image of the large field, many visual units that have received too few quanta to respond at all, the situation remains relatively simple. This is because the actual levels of stimulation of the other units, which have received the minimum or more than the minimum numbers of quanta required, are irrelevant. What matters in the present theory is whether or not units are stimulated, not their degree of stimulation.

It will be noted that at this stage we are dealing, not merely with the distribution of the acting quanta among the receptors at the extreme periphery of the nervous visual system, but with the problem of the functional organization of the human retina and of the visual system. On this problem, relatively little information, at least of a quantitative nature, is available from independent sources such as anatomical studies. The problem is a complex one; for instance, many experiments on spatial summation at threshold have been made with the use of concentric test-fields of increasing diameter, instead of two light spots of increasing separation. The results of these two types of

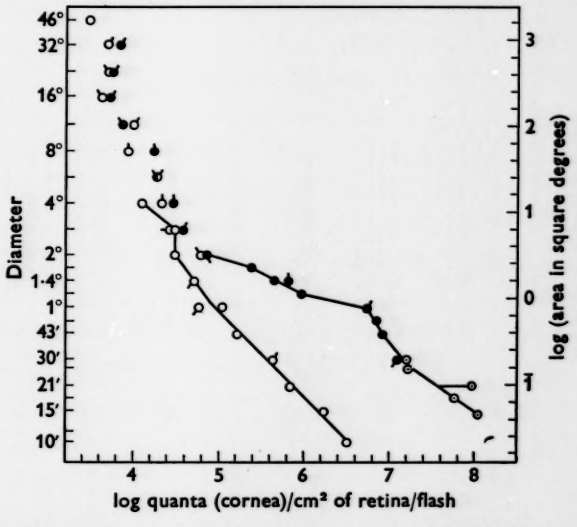


FIGURE 10 – *Results of measurements made by the method of figure* 8 *with black disks of varying diameters (black spots), and of absolute-threshold measurements made under similar conditions with luminous fields presented against a zero-intensity field (rings). All test objects centred* 20° *from the fixation point. Light flashes* 0·0026 *second in duration. Spectral band centred near* 507 mμ. *Black disks smaller than* 15′ *of arc in diameter could not be seen even at higher intensities. The straight line at* 45° *drawn through the empty circles for the smaller field diameters corresponds to complete summation, that is, to a constant total amount of light entering the eye at threshold, independently of the area of the luminous test field. (From P. E. Hallett,* [20].)

experiment do not agree quantitatively with regard to the spatial extent of the physiological summation, which seems to be greater in the concentric-field experiments; this fact still requires explanation. On the other hand, the quantum theory of light has helped to remove some puzzles concerning the working of the retina. Thus in the past, mere probability summation was confused with partial physiological summation, and this led to the erroneous view that physiological interaction occurred to a limited degree over the whole of the dark-adapted retina under absolute-threshold conditions.

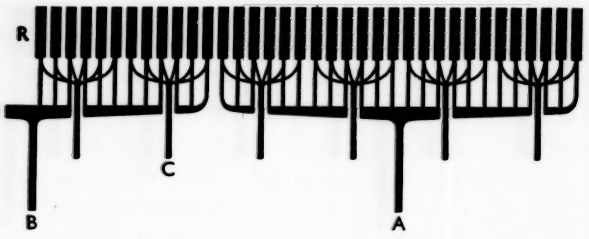


FIGURE 11 – *Schematic diagram indicating one possible type of nervous connection by which the same retinal receptors* R *would be organized into functional units of different sizes. The smaller units* C *should have a higher threshold, in terms of the illumination per unit area of the retina, than the large units* A *or* B, *since each of them could collect light only from a smaller area. (After Pirenne and Denton,* [21].)

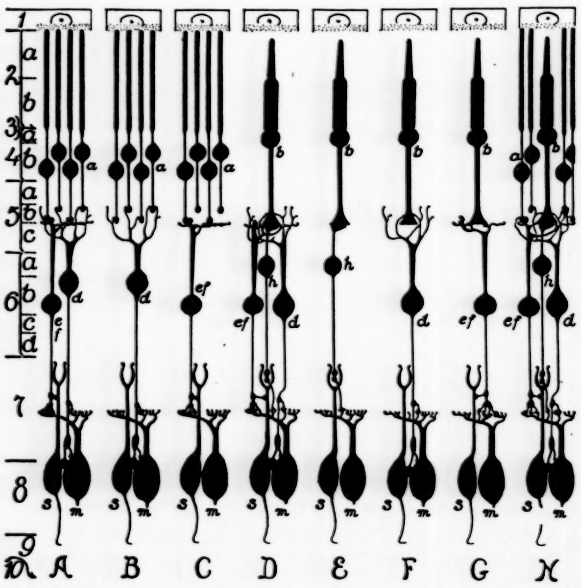


FIGURE 12 – *Types of connections observed anatomically between the rod and cone receptors and the fibres of the optic nerve. A, B, and C show how one optic nerve can be connected to several rods. The actual number of rods so connected may number tens of thousands. (From S. Polyak,* [22], *copyright* (1941) *of the University of Chicago Press.)*

THEORY OF THE ACUITY EXPERIMENTS

Because of quantum fluctuations, the distribution of the visual units stimulated by light will be random and vary from one flash to another; it will therefore resemble the distribution of the dots in figure 9, but here the size of the visual units is an important factor. These units probably overlap in the human retina, but there is little information on this point. It is clear that considerable complications in the quantitative aspect of the theory must be expected when the light flash becomes intense enough to cause stimulation of a large proportion of the units, and when the area of the image of the test object becomes about the same as the area of a unit, especially as the image is then likely to cover only parts of some of the units. If the visual units stimulated are few and far between, however, these complications do not arise. Under such conditions, and even at somewhat higher luminances, it is found that the threshold luminance, expressed as light intensity per unit area, is about the same for the detection of a black disk in the present experiments as for the detection of a luminous circular field having the same area—the latter arrangement corresponding, as it were, to the negative of the former, in the sense used in photography. At higher luminances, the curves for the threshold of visibility of a black disk and of a luminous circle diverge, showing that further complications begin to arise (figure 10).

The coincidence of the two curves at the lower luminances supports the theory, since it shows that the mean energy of the light that the black disk prevents from reaching the retina is the same as that of the light required to detect a luminous disk of the same area. Furthermore, in the curve relating to the black-disk experiments, a break occurs at a disk-area that is of the same order as the area within which complete summation is found to occur in the case of the luminous disk presented against a zero-intensity field (figure 10). This area probably gives the order of magnitude of the area of the visual units involved.

There are various reasons to suppose that the same retinal rods are arranged to form both large and small units, but at low luminance levels only the large units are in practice able to respond, because extremely few of the small units then receive enough quanta to be stimulated (figure 11). One would therefore expect to find a progressive recruiting of the large units as the flash luminance is increased and then, when they are all excited, a further progressive recruiting of the smaller units.

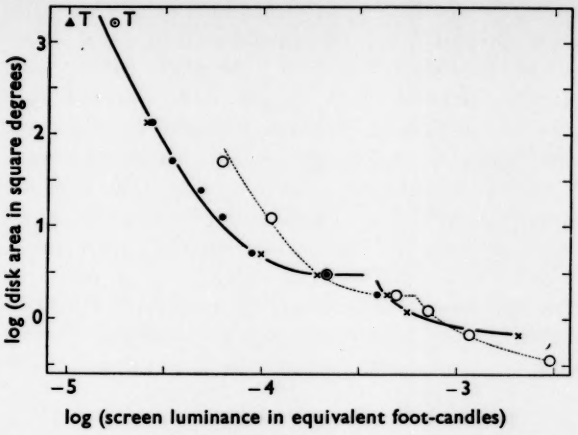


FIGURE 13 – *Results of experiments similar to those of the black disk experiments of figure* 10, *for another subject. The results are given for two retinal positions,* 20° (*solid curves*) *and* 10° (*dotted curves*) *from the fixation point. The experiments were made with white light in flashes of* 0·03 *second duration.* (*From Pirenne,* [23].)

Eventually the smallest units of the retinal periphery, consisting of cones only, or of both rods and cones, must become excited.

Figure 13 shows the results of earlier experiments performed with a different subject and under slightly different conditions. These experiments relate mainly to the limits of visibility of black disks, and cover a smaller range of disk diameters. They show that the results differ considerably at 20° and at 10° from the fovea; this was also observed in other experiments made by the subject of figure 10. Even within the periphery of the retina, the properties of the visual system vary from one position to another. This is a state of affairs observed in many types of experiments; it is certainly of practical importance so far as the use of the eyes is concerned, but it does greatly complicate the testing of visual theories.

Although the results of figures 10 and 13 are in general agreement, there are differences between them, which are probably due to differences in the experimental conditions, or perhaps to differences between the eyes of the two human subjects. Experiments of this kind, in spite of their apparent simplicity, entail a considerable effort of psychological concentration on the part of the subject. They are also time-consuming, since each experimental point is given by a statistical determination based on results corresponding to many flashes. After a couple of hours, the subject's fatigue tends to lead to irregularities in his responses. Consequently, the number of points that can be determined at one sitting is limited. Furthermore, on account of day-to-day variations in the same subject, it is not always easy to co-ordinate the results obtained at different sittings. Finally, special precautions must be taken to prevent the experimenter from unwittingly influencing the subject's results. The well-known story of the alleged discovery of the 'N-Rays' has demonstrated the effect psychological suggestion can have on careful and completely honest visual observers working under difficult conditions [24].

The above results constitute no more than one step in the study of visual acuity, and they have been discussed here because of the importance of this visual function.

SUBJECTIVE APPEARANCES

In absolute-threshold experiments made with a fairly large luminous field whose presence must be detected against a zero-intensity field, it is tempting to assume that the field is just visible when only one of the visual units covered by its retinal image is stimulated. But though this assumption is often made, there is no proof that it is true. The 'two-quanta theory' developed by H. A. van der Velden [12] and M. A. Bouman [30] is based on the assumption that the condition for seeing is that one visual unit be acted upon by at least two quanta—the number two being fixed, according to the theory. On the other hand, a more involved theory of the absolute threshold has been put forward by Baumgardt [25], according to which seeing a test field depends on more complicated 'coincidence' arrangements than the above hypothesis of physiologically independent visual units.

The subjective appearance of the field is of little immediate use in testing the above assumption. It might be thought that when only one, or two or three, units are excited within the area corresponding to a fairly large test field, the subject should see one, two, or three discrete patches of light in definite positions. But reasonings of this kind are notoriously unreliable, because, for one thing, there would be no practical or biological purpose in our 'seeing' the actual stimulation pattern in our retina. Some experiments were made in which a subject was presented (*a*) with a large circular field exposed in flashes at a fixed luminance near threshold; (*b*) with the top half only; or (*c*) with the bottom half only, of the field [26]. The experimental conditions, including the luminance of the fields, remained the same, and the actual sequence of the presentation of (*a*), (*b*), and (*c*) was unknown to the subject.

The subject then made confusions between the

half-fields (*b*) or (*c*) and the whole field (*a*); he mistook, at times, a half-field for the whole field. This is in keeping with the theory that only a few, sometimes perhaps only one, of the visual units covered by the field are stimulated at threshold. Then it must be expected that the stimulation caused by the whole field will sometimes be confined, say, to the top half only, and this will explain the confusions that occurred.

It would be more straightforward to find out whether the whole field can be confused with a single small test-field covering an area about the size of the retinal area inside which physiological summation is known to take place, or with random arrays of a few such areas. But the same luminance must be used for the whole field as for the small fields. If a higher luminance is used for the small fields, the stimulation of the visual units concerned is likely to be higher in the latter, and this may introduce a significant difference between the two modes of stimulation, for it is clear that a few really bright spots of light are not confused with a large extended field. Here the difficulty arises that if only one or a few small fields are used at a luminance near the threshold luminance (50 per cent frequency of seeing) of the large field, the frequency with which the subject's threshold is reached in the case of the small fields becomes so low that in practice he hardly ever has the opportunity of comparing their appearance with that of the large field. This effect on the threshold luminance is readily explained on the basis of probability summation.

In the case of the black-disk experiments also, the luminous field is not seen as a pattern of light spots; the unstimulated regions predicted by the theory are not perceived subjectively. At luminances well above that at which the large luminous field itself can just be seen, but adjusted to a luminance below that necessary to see the black-disk, the disk of course remains entirely invisible. Yet subjectively the field may then appear quite bright and very evenly illuminated. This will be so, for example, for a black disk 1° in diameter, twice the angular diameter of the sun; judging from the subjective appearance of the field, a subject unused to the experiments may find it hard to believe that he failed to see such a conspicuous object in the experiments.

Apart from the difficulty of maintaining attention, one of the problems confronting the human subject in absolute-threshold experiments is that of avoiding confusion of the test field with the ever-changing luminous clouds, the 'phosphenes', are subjectively seen in total darkness. Some subjects are much less troubled by phosphenes than others, but darkness as a rule never looks completely dark or black to the dark-adapted eye. The visual field is filled with *Eigenlicht*, the eye's own light. No doubt such effects are due to some form of spontaneous activity in the nervous system, but their exact origin is uncertain. They may, or may not, be related to the 'dark noise' which, according to some theories [27], exists in the human visual system. This dark noise would consist of spontaneous nervous activity having characteristics indistinguishable from those of the activity caused by the action of light on the retina. It might, for instance, be due to the thermal activation of rhodopsin molecules in the rods, on the assumption that such activation causes the same effects as the activation due to actual light absorption. There is no reason to suppose that such dark noise would necessarily lead to luminous sensations in complete darkness, as these would be of no apparent biological use. But since, by definition, the dark noise would be of the same kind as the nervous stimulation caused by light, it would necessarily interfere with the detection of actual light stimuli if the noise were of sufficient magnitude. This might explain why more than one acting quantum is necessary for vision—even though the strange effects which would then occur with regard to summation might also make one-quantum vision biologically useless. At the present time there does not seem to exist any convincing proof of either the presence or absence of such dark noise at a significant intensity in the human visual system.

CONCLUSION

Most of the experiments discussed above were made with brief light-flashes. With regard to visual acuity, the processes involved in ordinary vision are probably much more complicated. Then the retina becomes adapted to the prevailing illumination, and the eye moves continually, making fixation pauses of the order of 0·1–0·2 second in the course of its jerky displacements. Under these conditions it is probable that the perception of dark objects involves 'off' responses in the visual system; that is, nervous responses that occur when the light is cut off from the retina. Such 'off' discharges of impulses have been observed in the case of electrophysiological preparations of vertebrate eyes, for instance in optic-nerve fibres of the frog [28]. They may occur in the human retina when, as a result of a displacement of

the eye, the image of a dark object suddenly falls on a previously illuminated area. In this connection, it must be remembered that there is much evidence indicating that stimulation of the retina under steady conditions leads to very inefficient vision, or to complete loss of vision, after a few seconds. Accordingly, it may be surmised that, when the eye looks at a large evenly illuminated field, the central part of this field soon becomes visually ineffective, since it is only near the edges of the large field that eye movements cause changes in the illumination reaching the retina. If so, the only means of affecting the retina when the eye tries to detect a small black object in the centre of a large field would be displacements of the dark image of the object on the retina. There are reasons for believing that over the whole range of rod vision, and over a great deal of the range of cone vision, it is the cutting off of the retinal illumination by this dark image that is the determining factor for the response of the visual system at the limit of visibility of the black object [29].

The experiments discussed previously can do no more than indicate the ways in which the conception of quanta of light can be applied to the study of the visual process. The quantum theory does not solve the problems of the physiology of vision, but it allows many of them to be studied in a more concrete and exact manner, and so helps to reveal their complexity.

ACKNOWLEDGMENTS

Some of the experimental work reported in this article has been supported by the Medical Research Council, the Nuffield Foundation, and the United States Public Health Service (Grant No. B-1810 from the National Institute for Neurological Diseases and Blindness).

REFERENCES

[1] STILES, W. S. 'The physical interpretation of the spectral sensitivity curve of the eye'. In 'Transactions of the Optical Convention of the Worshipful Company of Spectacle Makers'. Pp. 97–107. Spectacle Makers' Company, London. 1948.
[2] SCHULTZE, M. *Arch. Mikr. Anat.*, **2**, 175, 1866.
[3] STILES, W. S. *Proc. roy. Soc., B*, **127**, 64, 1939.
[4] RUSHTON, W. A. H. *Progr. Biophys.*, **9**, 239, 1959.
[5] STILES, W. S. *Ann. R. Coll. Surg. Engl.*, 1961. (In press.)
[6] STILES, W. S. and CRAWFORD, B. H. *Proc. roy. Soc., B.*, **112**, 428, 1933.
[7] STILES, W. S. *Sci. Progr.*, **33**, 676, 1939.
[8] FLAMANT, F. and STILES, W. S. *J. Physiol.*, **107**, 187, 1948.
[9] PIRENNE, M. H., MARRIOTT, F. H. C., and O'DOHERTY, E. F. *Spec. Rep. Ser. med. Res. Coun. Lond.*, No. 294, 1957.
[10] HECHT, S., SHLAER, S., and PIRENNE, M. H. *J. gen. Physiol.*, **25**, 819, 1942.
[11] BRUMBERG, E. M., VAVILOV, S. I., and SVERDLOV, Z. M. *J. Phys. Moscow*, **7**, 1, 1943.
[12] VAN DER VELDEN, H. A. *Physica*, **11**, 179, 1944.
[13] PIRENNE, M. H. *Progr. Biophys.*, **2**, 192, 1951.
[14] PIRENNE, M. H. and MARRIOTT, F. H. C. In 'Psychology: A Study of a Science. Study 1. Conceptual and Systematic. Vol. 1. Sensory, Perceptual, and Physiological Formulations'; S. Koch (Ed.). Pp. 288–361. McGraw-Hill, New York. 1959.
[15] LUDVIGH, E. and MCCARTHY, E. F. *Arch. Ophthal. N.Y.*, **20**, 37, 1938.
[16] RUSHTON, W. A. H. *J. Physiol.*, **134**, 30, 1956.
[17] BAUMGARDT, E. 'Les théories photochimiques classiques et quantiques de la vision et l'inhibition nerveuse en vision liminaire'. Editions de la Revue d'Optique, Paris. 1950.
[18] Ref. [11], p. 56.
[19] PIRENNE, M. H. *Nature, Lond.*, **152**, 698, 1943.
[20] HALLETT, P. E. *Proc. 3rd Int. Congr. Photobiol.* (Finsen Memorial Congress, Copenhagen). (In press.)
[21] PIRENNE, M. H. and DENTON, E. J. *Nature, Lond.*, **170**, 1039, 1952.
[22] POLYAK, S. 'The Retina'. University of Chicago Press, Chicago. 1941.
[23] PIRENNE, M. H. *Proc. Camb. phil. Soc.*, **42**, 78, 1946.
[24] PIÉRON, H. In 'Mechanisms of Colour Discrimination'. P. 15. Pergamon Press, London. 1960.
[25] BAUMGARDT, E. *Proc. Int. Illum. Comm. XIV* Session, Brussels, Vol. A, p. 78, 1959.
[26] PIRENNE, M. H. *J. Physiol.*, **107**, 47*P*, 1948.
[27] BARLOW, H. B. *Ibid.*, **141**, 337, 1958.
[28] HARTLINE, H. K. *Amer. J. Physiol.*, **121**, 400, 1938.
[29] PIRENNE, M. H. *Ann. N.Y. Acad. Sci.*, **74**, 377, 1958. MARRIOTT, F. H. C. and PIRENNE, M. H. *J. Physiol.*, **152**, 24*P*, 1960.
[30] BOUMAN, M. A. *Studium gen.*, **13**, 491, 1960.
[31] CRAWFORD, B. H. *Proc. Phys. Soc. Lond., B*, **62**, 321, 1949.
[32] *Proc. Int. Illum. Comm.*, **3**, 32, 1951.
[33] CRESCITELLI, F. and DARTNALL, H. J. A. *Nature, Lond.*, **172**, 195, 1953.
[34] SHLAER, S. *J. gen. Physiol.*, **21**, 165, 1937.

GENERAL READING

PIRENNE, M. H. 'Vision and the Eye'. Chapman and Hall Ltd, London. 1948.
Idem. Biol. Rev., **31**, 194, 1956.
MARRIOTT, F. H. C. *Symp. zool. Soc. Lond.*, No. 3, 67, 1960.
BRINDLEY, G. S. 'Physiology of the Retina and Visual Pathway'. Edward Arnold, London. 1960.
PIRENNE, M. H. and MARRIOTT, F. H. C. 'Visual Functions in Man'. In 'The Eye', H. Davson (Editor), Academic Press, London and New York. (In press.)
National Physical Laboratory Symposium No. 8, 'Visual Problems of Colour'. 2 vols. Her Majesty's Stationery Office, London. 1958.

Sir James Hall, Bt. (1761–1832)

V. A. EYLES

Sir James Hall was a friend both of Lavoisier and of James Hutton, the geologist who first advanced the modern theory of the origin of the igneous rocks. Hall devised a number of chemical experiments to test the validity of the theory by attempting to reproduce subterranean temperatures and pressures; he was the first geologist to use such methods successfully. The experiments, and the results they achieved, are described in this article.

Sir James Hall (figure 1), who was born on 17th January 1761, was the first person successfully to apply the experimental techniques of the chemist to geological problems. The experiments were carried out between 1798 and 1804, when the use of such methods by geologists was almost unknown. They were made to test the theories put forward by Hall's friend, James Hutton, whose 'Theory of the Earth' was first announced in 1785 [1].

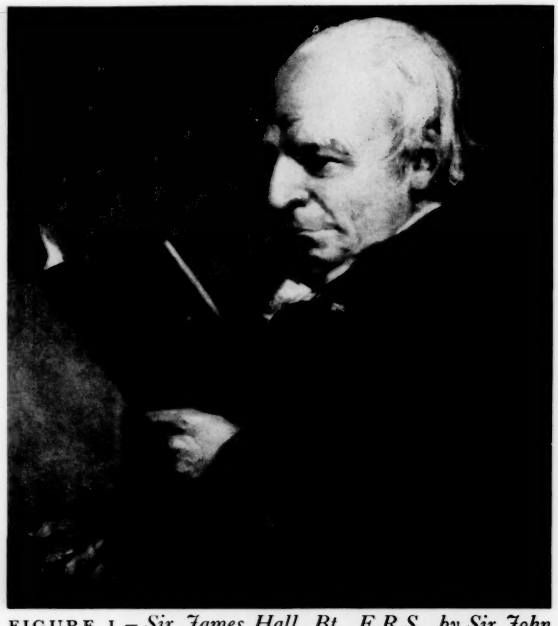


FIGURE 1 – *Sir James Hall, Bt., F.R.S., by Sir John J. Watson-Gordon, R.A.*

A short account of Hall's career will serve to show how he became involved in Hutton's problems. He was a descendant of an old Scottish family established in East Lothian. In his twelfth year he was sent to a school near London, and at the age of fifteen he succeeded to the baronetcy. Sir John Pringle, at that time President of the Royal Society, was his great-uncle, and he took young Hall to dine at the Royal Society Club, a dining club within the Royal Society, in London. This may have stimulated Hall's interest in science; but it is probable that he had already met Hutton, who was a neighbour and friend of his father. In 1777 Hall entered Christ's College, Cambridge, where he remained in residence for two years without graduating. After leaving Cambridge he studied first in France and then in Geneva. He returned to Scotland for his coming-of-age in 1782. In 1783 he went abroad again, for about three years, spent travelling in Europe. He was now actively interested in geology—during this tour he visited Vesuvius and the volcanoes in Sicily and the Lipari Islands. In Sicily, he was accompanied by a distinguished French geologist, Déodat Dolomieu. Hall made a careful examination of Vesuvius and the adjacent Monte Somma, and the sketches and notes he made at the time proved useful when he came to study Scottish geology. During 1786, he spent several months in Paris, and visited Lavoisier frequently. As a result, he became a convert to the latter's views on chemistry.

He returned to Edinburgh in November 1786 to marry Lady Helen Douglas. Early in 1788 he addressed the Royal Society of Edinburgh on 'Lavoisier's New Theory of Chemistry'. This discourse, occupying three meetings, roused both interest and opposition, for two subsequent meetings were devoted to an exposition of the phlogiston theory by Hutton, and by Hall's reply to Hutton. Hall's address was not published, but it was mentioned by Robert Kerr, the translator of Lavoisier's *Traité élémentaire de chimie*, in a letter written to Lavoisier on 21st January 1791. Kerr wrote that Hall had given the Royal Society of Edinburgh a very detailed account of Lavoisier's doctrine and had converted some people, including Kerr himself, to the antiphlogistic theory [2]. This letter shows that Hall was one of the earliest supporters of Lavoisier's doctrines in Britain. He had been deeply interested in chemistry since his

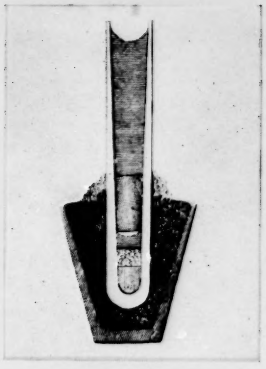


FIGURE 2 – *Porcelain tubes as used in the second stage of compression experiments.*

FIGURE 3 – *Original crucible, used by Hall in his 'whinstone' experiments, showing the glass formed by rapidly cooling a melt of igneous rock (height about 6 cm).*

FIGURE 4 – *A similar crucible, showing the crystalline mass obtained by slow cooling.*

FIGURE 5 – *Drilled iron bar with plug.* (7·5 *cm long.*)

FIGURE 6 – *Earliest apparatus used by Hall in his compression experiments.*

FIGURE 7 – *Drilled bar* (37 *cm long*).

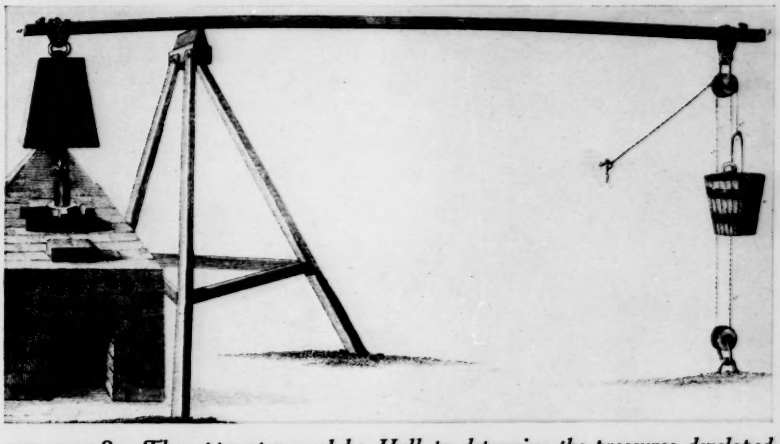


FIGURE 8 – *The apparatus used by Hall to determine the pressures developed during his experiments. A very heavy iron weight rests on the open end of a tube filled with fusible metal, and the reading is taken when the expansive forces of the gases generated start to lift the weight. Exact adjustments are made by counterpoise weights placed in the bucket at the end of the beam.*

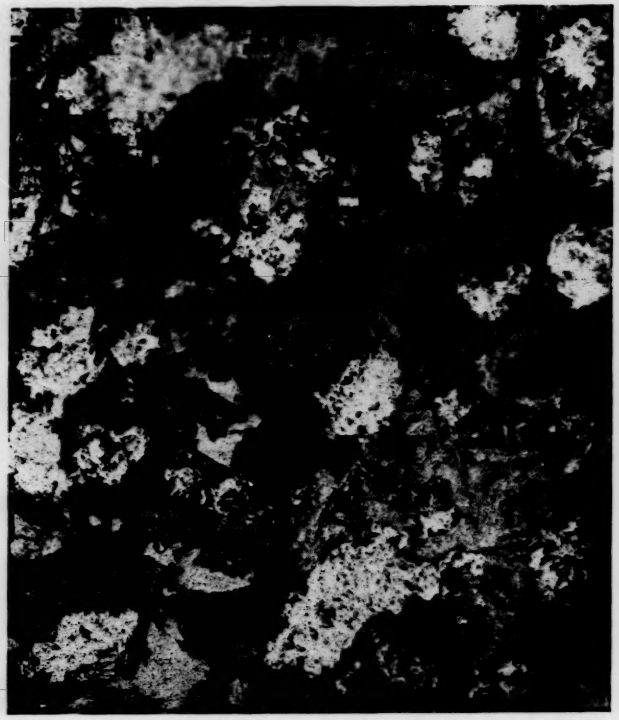


FIGURE 9 – *Thin slice of one of the 'whinstones' used by Hall, before melting, showing typical crystalline texture.*

FIGURE 10 – *Thin slice of slowly cooled igneous melt, showing crystallites in a glassy base.*

FIGURE 11 – *Another slowly cooled igneous melt. The texture is not that of a typical unmelted whinstone (see figure 9). (Figures 9–11 highly magnified.)*

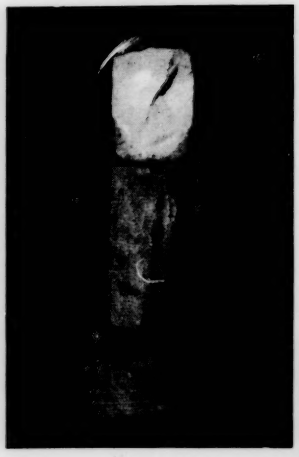


FIGURE 12 – *A specimen of 'marble' produced in one of Hall's experiments.*

FIGURE 13 – *Porcelain tube, sectioned, after an experiment.*

student days, and had attended the lectures of Joseph Black in Edinburgh.

It was at this time, in 1788, that the first full account of Hutton's 'Theory' was published. Hall's first reaction was unfavourable. He said, later, that his first inclination was to reject the theory, and that he would have done so had he not known the author, whose lively, perspicacious conversation was strikingly contrasted to the obscurity of his published writings. It was only after three years of almost daily conversation that Hall began to accept Hutton's fundamental principles. The change in opinion was a fortunate one, for it led Hall to his classic experiments.

Some knowledge of the state of geology at that time is necessary for an appreciation of the experiments. Many geological facts had been accumulated—by observation of surface rocks, by mining operations, and by the study of minerals; what had hitherto been lacking was a master theory that would co-ordinate the known facts and explain their significance. Much was already known about the sedimentary rocks—the limestones, shales, and sandstones—but there was another class of rocks, including the crystalline granites, basalts, and porphyries, about whose origin the geologists were almost completely ignorant. These are now called 'igneous' rocks. We know today that they originated in the Earth's interior, and that they have consolidated from a molten condition. The widely held view in Hutton's time was that, like the sediments, they had been laid down in water, or had crystallized from solution. It is remarkable, considering how much was then known about the solubility of chemical compounds, that this view should have been held, for the rocks are mainly composed of virtually insoluble silicate-minerals, such as quartz and felspar. A single exception to the general view was provided by lavas, a term then restricted to rocks that had obviously flowed from active or recently extinct volcanoes; most geologists thought that this phenomenon was purely local. They believed that many basalts in Europe and elsewhere, now known to be lavas, were aqueous deposits. This view continued to be held by some prominent geologists until the early years of the nineteenth century.

The first correct and comprehensive explanation of the origin of the igneous rocks was put forward by Hutton. He supposed that in the deeper parts of the Earth's crust all rocks were fused, and that great quantities of molten rock had moved up into the upper sedimentary parts of the crust. Some poured out on the surface as lavas and some consolidated underground. The latter formed 'subterraneous lavas' (flat sheets), vertical dykes (veins), or large irregular masses, such as the granites of Scotland. This was a major advance in geological thought, and those who accepted Hutton's views on the origin of igneous rocks came to be known as 'Plutonists', in opposition to the 'Neptunists', who held the earlier views.

Hutton's theory was much criticized on particular points, and we are especially concerned here with two criticisms that carried considerable weight at the time. They arose from ignorance of the chemistry of rock melts. This is not surprising, for the subject is very complex.

The first criticism concerned the fusion of igneous rocks. It was known that some of the commoner igneous rocks, when melted in a crucible and allowed to cool, formed a glass, not a crystalline rock. The Neptunists claimed that, if the rocks that Hutton supposed once to have been molten really had consolidated from such a condition, they would now be found as glasses and not as crystalline rocks.

The second criticism arose from the fact that limestone, or carbonate of lime, when heated, decomposes to quicklime, giving off carbon dioxide. This was a fact particularly well known to Edinburgh scientists, for it had been established by Black. Hutton now claimed that rocks forming land surfaces, including limestones, had, in the past, been raised to the surface after deep burial in subterranean regions where very high temperatures prevailed. Why, the critics asked, had the limestones not been decomposed?

Hutton's answer to both objections was that the conditions existing deep in the Earth's crust were very different from those on the surface; he thought it impossible to imitate them. He appears to have feared that experimental attempts to meet criticisms might do more harm than good if the experiments were unsuccessful. He later issued a warning against those who 'judge of the great operations of the mineral kingdom from having kindled a fire and looked into the bottom of a little crucible'.

Hall was, of course, aware of Hutton's views, but it seems that he still thought it worth while attempting to test the theory experimentally. He carried out a few preliminary experiments in 1790, but discontinued them out of respect for his friend's opinion. Hutton died in 1797 and Hall resumed his experiments.

His first were his 'Experiments on Whinstone

and Lava' [3]. He hoped to show that a melt of igneous rock, if cooled slowly, would produce a crystalline rock, and not a glass. He used the Scottish term 'whinstone' for rocks that, in Hutton's view, were igneous intrusions, and the term 'lava' in its then restricted sense, for rocks known to have been extruded by a volcano. It had already been demonstrated by James Keir [4] that crystallization could take place in ordinary glass if cooled slowly; and Hall had been reminded of the fact more recently by seeing, in a local glass works, the result of slow cooling of a melt of green bottle glass, which had become a stony mass. He believed, therefore, that his chances of success were good. He used eight different 'whinstones' and seven lavas, some of which he had brought back himself from Vesuvius and Sicily. Each specimen was first heated in the reverberatory furnace of an iron foundry until thoroughly molten, and then cooled rapidly to a glass. Fragments of the glass so formed were then reheated until molten, and held for some time at a lower temperature in an open fire or in a muffle furnace before being allowed to congeal. Hall soon found that by regulating the rate of cooling he could produce a stony, more or less crystalline, mass. He discovered by trial and error how long to keep his melts in the muffle furnace—he mentions twelve hours for one experiment. The preliminary conversion to glass was a precaution against the criticism that he had not heated his specimens sufficiently to destroy their crystalline texture (figures 3 and 4).

It seems clear that in all Hall's experiments the final product showed some evidence of the presence of crystals. In some it was only incipient crystallization, in others he obtained rocks that may have resembled superficially those with which he had started. The only means he had for examining them was a hand lens; the use of thin rock-sections had not then been developed. Fortunately, some of his material is now in the Geological Museum, London. This has been examined microscopically, and, in one melt, crystals of felspar, augite, olivine, and iron ore were recognized, minerals characteristic of some of the rocks melted. In others, though crystallites or centres of spherulitic crystallization had developed, a glass matrix remained. Hall does not seem to have succeeded in producing a rock with a crystalline texture comparable to that of the rocks with which he started (figures 9–11). This is not surprising, for the time taken for naturally occurring igneous rocks, even lavas, to cool is very much more than that allowed by Hall.

Hall kept a check on the temperatures reached during his experiments by means of 'thermometric pieces', or pyrometers. These were invented by Josiah Wedgwood the elder for measuring temperatures too high to be recorded by any thermometer then available. Unfortunately, it is impossible to convert the figures on Wedgwood's scale to those on standard thermometers; but in some experiments his material was raised to temperatures exceeding the melting point of silver, about 950° C.

Hall's experiments on whinstones and lavas were interesting and valuable, but his next investigation, on 'The Effects of Compression in Modifying the Action of Heat' [5], was more important. He heated carbonate of lime to high temperatures and under high pressures to demonstrate that under these conditions it would not dissociate to quicklime and carbon dioxide. He also hoped to form crystalline marble, and even to melt the carbonate of lime. He carried out, between 1798 and 1805, more than 500 experiments. It was a process of trial and error which, as his experimental skill increased, led finally to success. Hall describes the methods and the results in much detail; it is possible only to summarize his account here.

He experimented with different kinds of calcium carbonate. He used, generally, powdered chalk; but he also worked with, for example, common limestone, marble, and chemically pure carbonate of lime. Hall realized that the carbonate of lime that occurs naturally is seldom pure. The amounts he used for each experiment varied from about half a gram to five grams.

His earliest attempts were made with a container with a lid fixed with screws, but this method was soon abandoned for a much simpler one. This used iron tubes of narrow bore, sealed at both ends. These were, in fact, gun-barrels, the breech being closed by welding in an iron plug. The carbonate, wrapped in paper, was inserted into the breech, and the rest of the barrel was then filled with baked and powdered clay. The muzzle was closed with another iron plug, and the end of the barrel containing the carbonate was put into a muffle furnace, the muzzle remaining outside, where it was kept cool with wet cloths (figure 6). One of the earliest results was the explosion of the furnace, but this did not deter Hall. He constantly modified his apparatus. He was not satisfied with the results obtained

using powdered clay to fill the barrel, and substituted a fusible metal that melted at the temperature of boiling water and could easily be removed after the completion of an experiment. This method was used for some time, until another difficulty arose. The fusible metal sometimes diffused through the iron barrel, from which it was extruded as fine wire. This, Hall thought, must have been the result of the pressure set up by expansion of the metal. He therefore inserted with the carbonate a small tube, open at one end, containing air, which, he thought, would act as a cushion to the expansion of the fusible metal. This method seems to have succeeded, once he had discovered, after destroying many gun barrels, the correct amount of air to use. He prevented contamination of the carbonate by placing it in small tubular containers, about 6 mm in diameter. He used glass first; but this proved too fusible, and he substituted porcelain. The containers were held in an iron frame on the end of a ramrod, so that they could easily be removed after an experiment (figure 6). He sometimes used lead instead of the fusible metal, and he also put powdered 'silex' (probably flint or quartz) into the porcelain container to prevent direct contact between carbonate and fusible metal. A major change was the substitution of porcelain tubes (about 34·5 cm long, having a bore of 6 mm and walls 5 mm thick), obtained from Wedgwood's Etruria works, for the gun-barrels previously used (figures 2 and 13). He also used silex as a filler, and left a space at the open end of the tube to be filled with melted borax to make the tube gas-tight, but this proved unsatisfactory. Hall suspected that the porcelain tubes were porous, and he used melted borax, with some success, as an internal glaze. Later, being dissatisfied with Wedgwood's tubes, he manufactured more satisfactory ones himself from a mixture of Cornish porcelain clay and 'Cornish stone' (possibly china clay). He varied his method of placing the tubes in the muffle furnace, finding it more satisfactory, when using borax as a seal, to insert them vertically (figure 2). The increased gas-tightness permitted higher pressures to be used.

By 1803 Hall had carried out 156 experiments, with variable results; in some, the loss of carbon dioxide was quite small, 2–3 per cent, or even less. At this stage he was quite justified in concluding that calcium carbonate heated under pressure was not significantly dissociated even at temperatures above 1000° C. It also seems probable that he had succeeded in converting powdered chalk into something closely resembling crystalline marble (figure 12). In one experiment, using 80 grains (about 5 grams) of powdered chalk, he produced a compacted semi-transparent mass, the interior of which had the characteristic rhomboidal cleavage of calcite. The loss in weight was only 3·5 grains, less than 5 per cent.

Hall still hoped to show that the calcium carbonate could be melted. He decided to increase both temperature and pressure, and decided that the porcelain tubes he had been using would not prove satisfactory. He tried gun-barrels again, but found them unsuitable too. So he decided to use iron bars, in which he drilled holes for the porcelain containers (figures 5, 7). He was able to obtain a supply of bars of an exceptionally high-grade iron, known as 'Old Sable' iron, which was imported from Russia. A friend, Dr Robert Kennedy, a skilled chemist, suggested he should place in the barrel, with the carbonate, small amounts of water, which, on vaporizing, would greatly increase the pressure developed. Hall adopted this suggestion, and even thought he could improve on it by using, in some experiments, ammonium carbonate, ammonium nitrate, gunpowder, or paper impregnated with nitre, but he found them no improvement on water. One result of these methods was an interaction between the carbonate and the clay container, forming one or more new minerals, which appear to have been rendered molten, or to have been redeposited after sublimation.

In 1804 Hall publicly claimed that he had succeeded in fusing the carbonate. It was suggested that this may have been the consequence of the presence of impurities, so he decided to use platinum as a container. He could obtain it only in the form of thin sheets, which he folded like a filter paper, to make a cup. He also obtained a supply of specially prepared, chemically pure, carbonate of lime.

As in his whinstone experiments, Hall kept a careful check on the temperatures reached by means of 'pyrometer pieces' of the type invented by Wedgwood. Although it is difficult to know exactly the temperatures reached, Hall, fortunately, noted that 22 on his scale corresponded to the melting point of silver, which allows a very rough guess to be made of the temperatures actually reached. These must often have considerably exceeded 1000° C.

Not content with recording temperature, Hall devised a means of determining the pressure developed during his experiments, based, he states,

on the means used by Count Rumford to measure the explosive force of gunpowder. The apparatus was crude by present-day standards (figure 8), but he could obtain approximate figures. The pressure was deduced from the weight lifted by the expansion of the gases generated in his compression tubes during an experiment. The maximum he recorded was 270 atmospheres; generally the figure was much less. This, he calculated, was equivalent to a column of sea water 8896 ft (2720 m) in height. He was, of course, interested in finding out the pressure exerted on rocks buried under a great depth of sea water or a thick cover of rock.

Hall accomplished most of his aims. He certainly established that calcium carbonate is not dissociated when heated under pressure, and from a geological point of view this was much the most important of his achievements. He also seems to have succeeded in turning powdered chalk or limestone into something very like a true crystalline marble, throwing light on the conditions of formation of certain naturally occurring marbles. Whether or not he really succeeded in melting calcium carbonate has been questioned. It is impossible to be sure from his descriptions, because, in those experiments in which he claims success, he does not mention having applied any test, such as effervescence with acid, which might have established the nature of the melt produced. Even with the resources of a modern laboratory it is not easy to melt calcium carbonate, and there is no unquestioned record of its having been done before 1912, when it was effected by H. E. Boeke, at a temperature of 1289° C and a pressure of 110 atmospheres [6]. It seems just possible that the necessary conditions may have been attained in some of Hall's experiments, but we shall never know for certain. What can be said, however, without fear of contradiction, is that what Hall did achieve was a very remarkable performance for an eighteenth-century amateur scientist, and of lasting value for geological science. He also deserves to rank as a distinguished pioneer in the history of physical chemistry.

Hall's results attracted much attention at the time, both in Great Britain and abroad, and they have often been quoted since. A French edition of his compression experiments was published in Geneva in 1807. It was translated by Marc-Auguste Pictet, a prominent Swiss scientist, who compared Hall with Lavoisier in a laudatory preface [7].

After the completion of the experiments described above, Hall continued to study geology and published several geological papers of interest. It is also worth remembering that his work shows him to have been not only an experimental geologist but a capable field geologist. He was elected a Fellow of the Royal Society of London in 1806, and in 1812 became the second President of the Royal Society of Edinburgh, an office he held until 1820. He died on 23rd June 1832 in Edinburgh, after some years of illness.

ACKNOWLEDGMENTS

Thanks are due to the Director of the Geological Survey and Museum, London, for permission to reproduce the illustrations of Hall's original apparatus, in the possession of the Museum; and to the Council of the Royal Society of Edinburgh for allowing reproduction of the portrait of Hall, which is in their possession, and the remaining illustrations taken from the accounts of his experiments in Volumes V and VI of the Society's 'Transactions'.

REFERENCES

[1] MacGregor, M. *Endeavour*, **6**, 109, 1947.
[2] McKie, D. *Notes roy. Soc. Lond.*, **7**, 13, 1949.
[3] Hall, Sir James. *Trans. roy. Soc. Edinb.*, **5**, 43, 1805.
[4] Keir, J. *Phil. Trans.*, **66**, 530, 1776.
[5] Hall, Sir James. *Trans. roy. Soc. Edinb.*, **6**, 71, 1812.
[6] Clarke, F. W. *Bull. 770 U.S. geol. Surv.* (fifth edition), 563, 1924.
[7] Pictet, M. A. 'Description d'une suite d'expériences qui montrent comment la compression peut modifier l'action de la chaleur', p. ix. Geneva. 1807.

Surface patterns in Boreal peatland

H. SJÖRS

The Boreal peatlands—that is, the bogs and fens in the zone around the North Pole where drained ground is mainly covered by coniferous forest—show very striking patterns. The origin of these patterns is not yet clear, and will become so only after extensive research. The present state of the study of possible origins is described and illustrated in this article.

It has been known for a long time that the northern bogs and fens show clearly defined patterns in which firm, dry areas alternate with soft, highly waterlogged ones. The origin of these striking patterns is still, however, unclear. There are a number of reasons for this. The areas are not always easily accessible on foot, as Linnaeus vividly recorded [1], and correlation of the work of different scientists has been hampered by differences of language. Aerial travel and aerial photography have, however, vigorously stimulated the studies—the areas can be more easily reached and the patterns more clearly seen—and a survey of the present state of research is now appropriate.

Analysis of local and regional distribution of the peatland patterns can serve as a starting point for a detailed study of their hydrotopography and hydrodynamics, but the problems are not purely geomorphological; an understanding of the state, development, and origin of the structures forming the patterns must be based on ecological research, starting from the present vegetation and its environmental conditions, and working backwards through detailed peat stratigraphy.

THE BOREAL ZONE

The Boreal zone is defined as the almost circumpolar zone in the north where drained grounds are covered by prevailingly coniferous forest [2]. Although patterned peatlands are not confined to this zone, they are far more common there than elsewhere. The Boreal forest, often referred to as Taiga, reaches from Trondheim to the Sea of Okhotsk, and from Alaska to Newfoundland (figure 13). Oceanic coniferous forests, such as the pinewoods of Scotland and westernmost Norway, and the Pacific forests, are not included, nor are other forests of non-Boreal conifers, of which there is a great variety. It is possible to subdivide the Boreal zone on a basis of the distribution of forest trees and other species, but a zonal subdivision is of greater general and ecological value. Apart from the transitional zone to the south, which in turn passes either into regions of deciduous forest or into steppes, the Boreal zone may conveniently be divided into main Boreal and sub-Arctic subzones. The former has been further subdivided, for instance by the Russians, into 'southern Taiga' and 'central Taiga' [3]. The sub-Arctic (the 'northern Taiga') is separated from the treeless Arctic by a usually narrow transition of woodland tundra. The zonation due to differences of altitude in mountainous parts of the Boreal zone roughly parallels zonation by latitude. However, the problem of equivalence is somewhat complicated by the presence of more or less subalpine birchwoods, particularly in Fennoscandia and Kamchatka.

The zones and subzones are closely related to the duration of the growth season, but they are little related to winter cold. All over the Boreal zone the winter temperatures are below freezing point, and seasonal soil frost is general. On the other hand, persistent strata of frozen soil, 'permafrost', are found only in the most continental parts of the Boreal zone, where the winter temperatures are exceedingly low. Some of the structures and patterns in peatland are evidently closely related to the type of soil frost, but others seem to be independent of freezing phenomena; the patterns in permafrost areas must be considered separately.

The distribution of vegetation shows no consistent relationship to that of permafrost [4]. Thus, in Europe and Labrador, discontinuous permafrost occurs in parts of the sub-Arctic. In other parts of Canada and in western Siberia some discontinuous permafrost is found as far south as the main Boreal subzone, but most of it is in the sub-Arctic; in the more continental parts, permafrost becomes continuous some distance south of the tree-line. In central and eastern Siberia, continuous and extremely deep permafrost is found throughout the Boreal zone, and even further south.

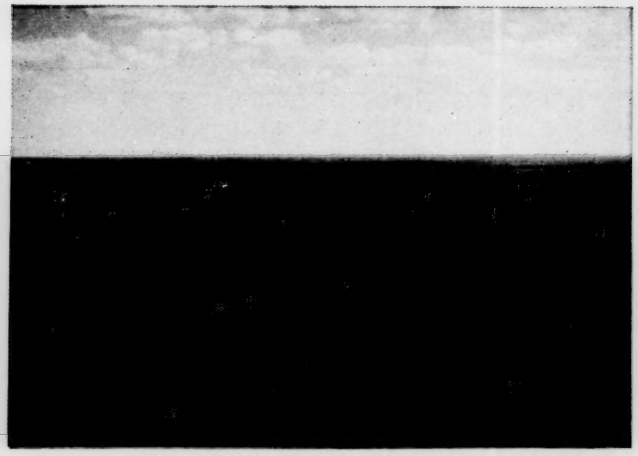


FIGURE 1 – *Expanse of minerotrophic peatland (fen). Small islands with a dense cover of black spruce* (Picea mariana). *All photographs on this page are from the Hudson Bay lowland in northern Ontario.*

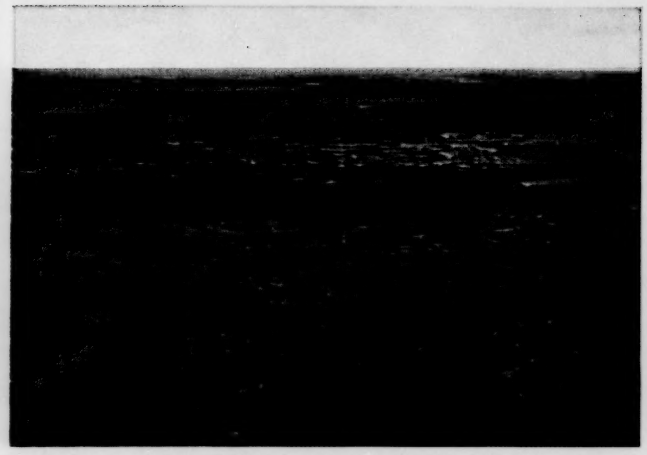


FIGURE 2 – *Ombrotrophic peatland (raised bogs) with numerous large bog-pools. A brook runs between two bogs (across the picture) and is expanded into a beaver pond (right).*

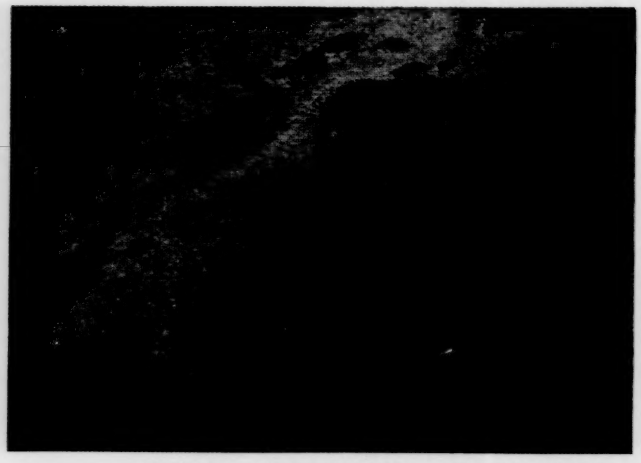


FIGURE 3 – *Left part ombrotrophic, colour mainly from* Sphagnum rubellum. *Right part minerotrophic, with pattern at right angles to slope, green colour partly from dwarf larch trees. Black spruce islands in centre.*

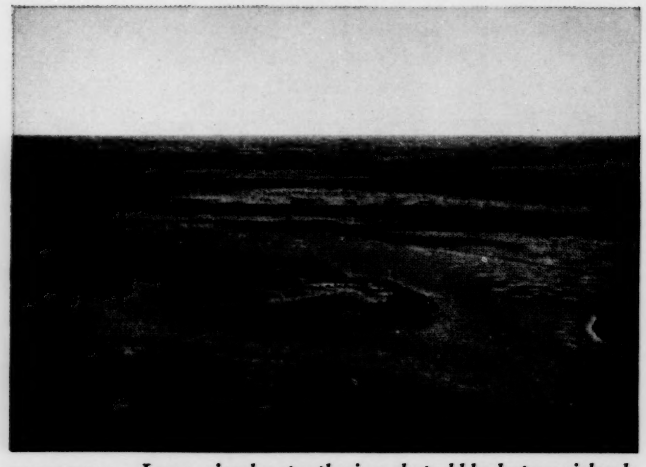


FIGURE 4 – *Large, circular, partly ring-shaped black spruce islands in a matrix of minerotrophic peatland largely covered by low larch trees* (Larix laricina).

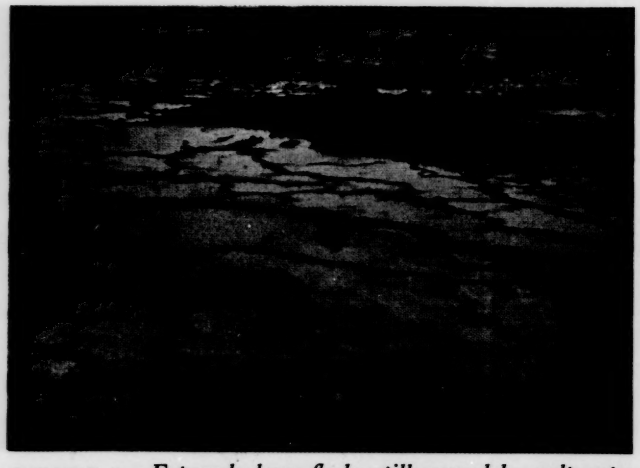


FIGURE 5 – *Extremely large flarks still covered by melt-water and separated by elevated ridges, partly with black spruce trees. This pattern is reminiscent of a rice-field.*

FIGURE 6 – *Lake probably formed from several coalescent pools. This type of pattern is found in almost flat areas which constitute water divides between rivers.*

FIGURE 7 – *Ombrotrophic peatland with bog-pools. Scots pine; dominants* Calluna, Betula nana, *and* Eriophorum vaginatum; *at water's edge* Scheuchzeria palustris *and* Carex limosa. *Tisjökölen, Central Swedish uplands.*

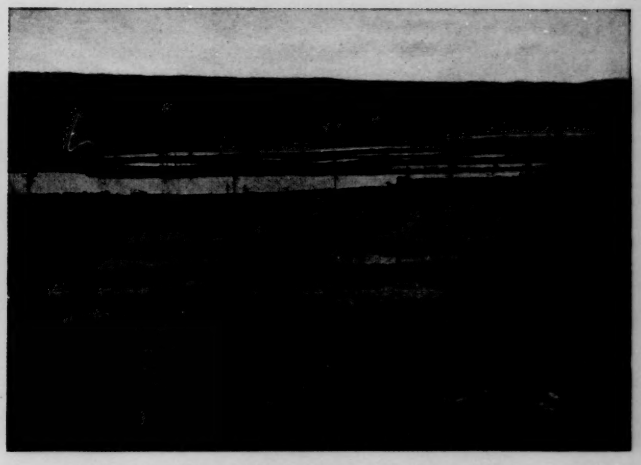


FIGURE 8 – *Ombrotrophic peatland with elevated ridges and a few bog-pools. Between ridges chiefly* Trichophorum caespitosum. *Foreground sandy, with* Cladonia *lichens and pine. Near Lofsdalen, Central Swedish uplands.*

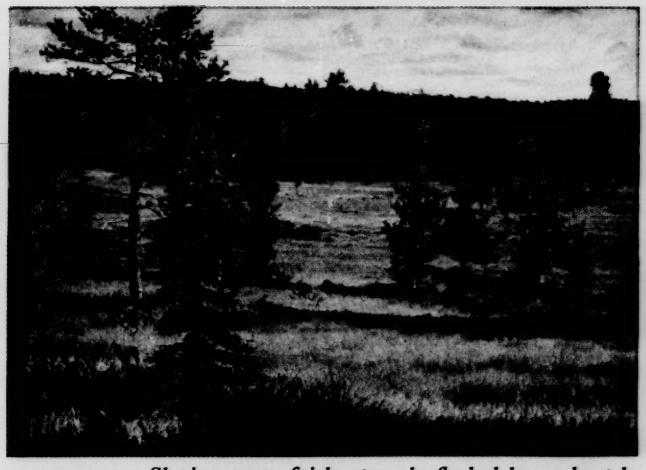


FIGURE 9 – *Sloping area, fairly strongly flushed by moderately minerotrophic water.* Molinia coerulea *in foreground. Sävsnäs, south part of Central Swedish uplands.*

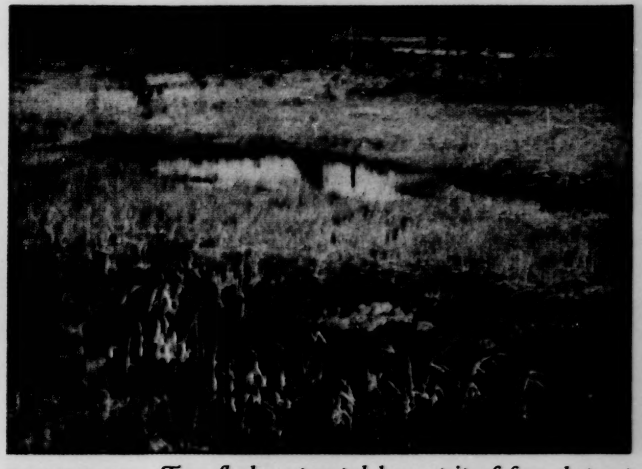


FIGURE 10 – *Two flarks separated by a strip of firm, but not much elevated, mat with* Carex *and* Molinia. *In the flarks* Carex rostrata, C. livida, *and* Rhynchospora alba. *Downhill of area shown in figure* 9.

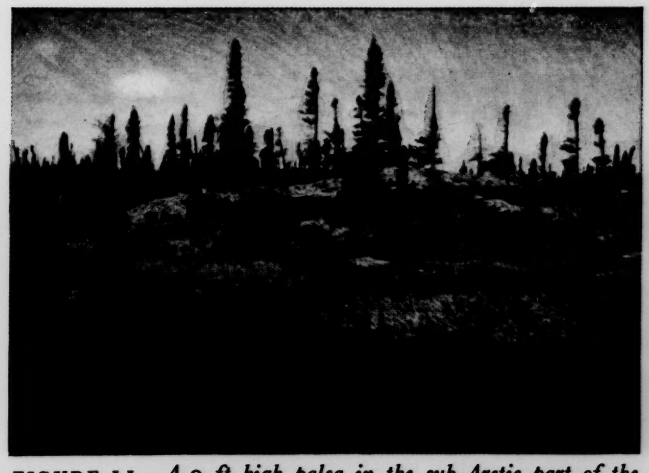


FIGURE 11 – *A* 9 *ft high palsa in the sub-Arctic part of the Hudson Bay lowland. Wind-worn trees of black spruce; carpets of lichens (largely* Cetraria nivalis) *indicative of lack of snow cover.*

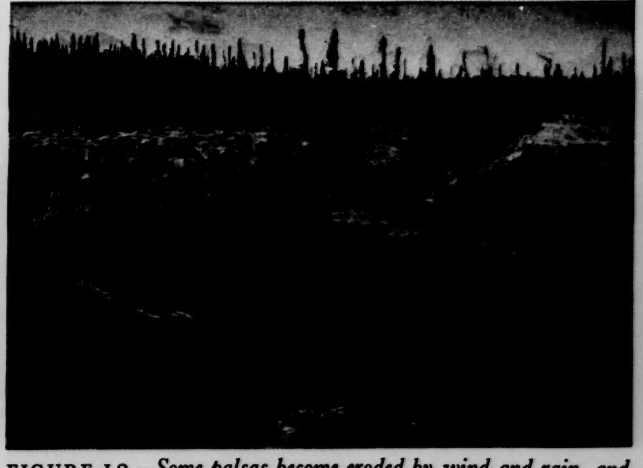


FIGURE 12 – *Some palsas become eroded by wind and rain, and then the ice core begins to melt. The pool is probably the site of a totally collapsed palsa. Same area as figure* 11.

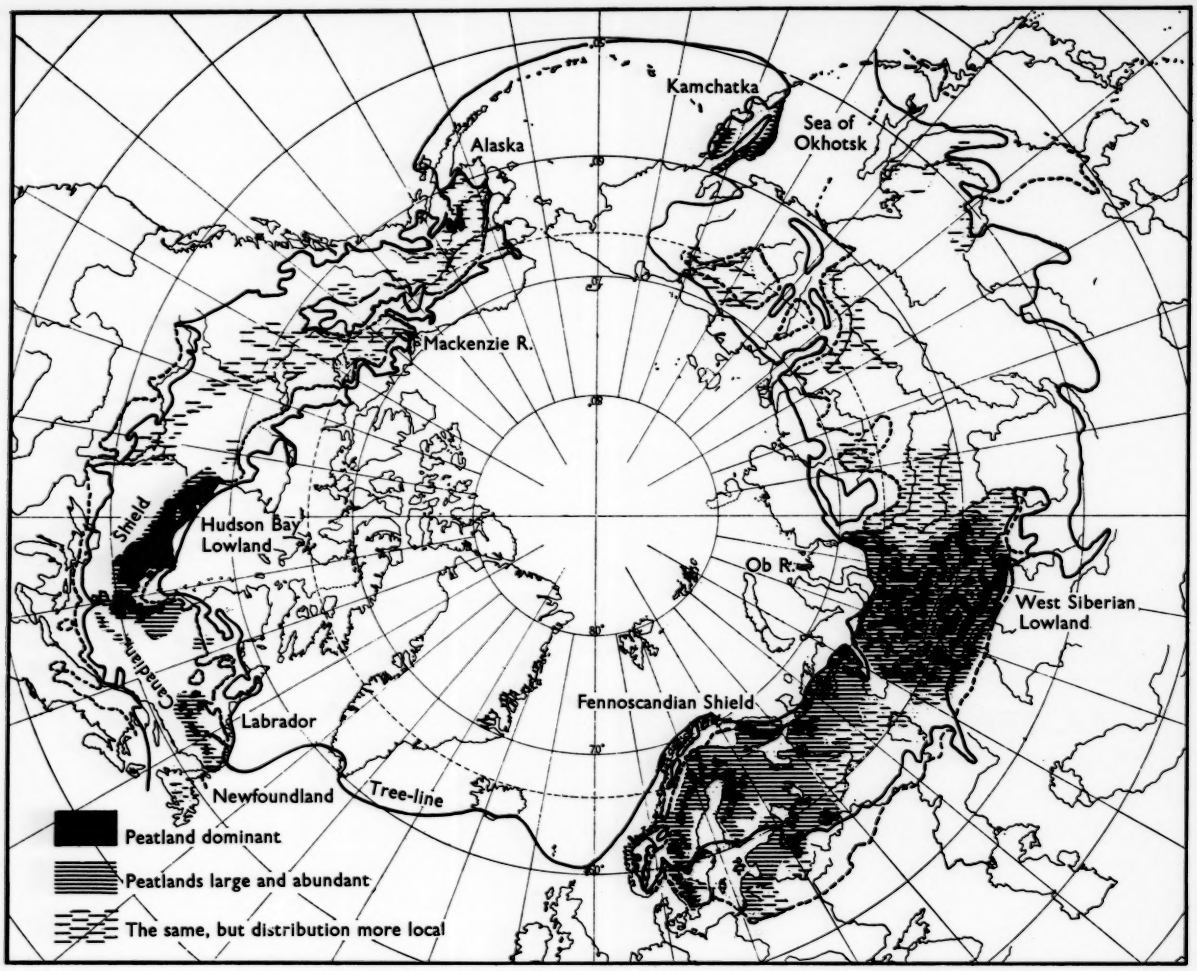


FIGURE 13 – *Area of abundant peatland in the Boreal zone. South of, or below, the tree-line, woodland tundra or subalpine areas extend to the northern broken line. South of the Boreal zone, transitional regions reach to the southern broken lines.*

The Boreal zone, like other biotic regions, is far from uniform in its vegetation. Boreal forests, some of which grow on peat, are less variable than more southern ones, but they still display a considerable diversity due not only to climate, but also to state of development, soil and water relations, and human influence. The most common types of unwooded areas are open bogs and fens, but their abundance is likely to be underestimated by an earthbound traveller.

PERCENTAGE OF PEATLAND

By far the greater part of the total peat area of the world is within the Boreal zone. However, because of variations in the relief of the land within this zone, the percentage covered by peat varies considerably. South of Hudson Bay there is a continuous lowland of about the same area as Great Britain. This lowland, with almost the only exception of lakes and river-beds, is completely covered by wet peat [5]. In other parts of Boreal America the peat deposits are less extensive, but there are considerable areas covered by peat, for instance in Alaska [6] and in Labrador [7].

The distribution of peat in Russia is uneven [8]. A large extent of western Siberia surrounding the Ob and Irtysh rivers is almost wholly covered by peatland, except along the rivers themselves and their tributaries. This is the largest peat area in the world. Some parts of Kamchatka and northern European Russia are also very rich in peat.

About one-third of Finland is covered by peat [9], the highest percentage for any country. In certain parts of Finland the percentage of land covered by peat reaches 60–70 per cent, and Sweden also has some areas where it covers more

than half the land. About 20 per cent of the Boreal zone in Sweden lies under peat. Individual peat areas in Sweden and Finland are rarely very large, because of the lack of unbroken plains, but a few are quite extensive. Thus, the largest continuous peat area in Lappland, the Sjaunja, covers approximately 150 square miles.

The type of peatland, as well as the area, depends on the topography of the land, and also on the climate. It is perhaps true that the relative importance of climate has been overestimated—it is now clear that highly different types of peatland can exist, in the Boreal zone and elsewhere, in areas with practically identical climates.

THE NUTRITION OF PEAT VEGETATION

Two sources of mineral nutrients are available to peat, the atmosphere and the soil. The former obviously benefits all peatland, chiefly by means of the precipitation and dust that fall on them. The terrestrial supply, on the contrary, is very variable, and some peatlands, or parts of them, do not receive any at all. Areas dependent exclusively on nutrients from above are termed 'ombrotrophic' [10], that is, 'nourished by rain.' The sites that, in addition, receive terrestrial mineral nutrition are termed 'minerotrophic,' indicating that nutrients are brought to the peat by water that has previously extracted them from a mineral soil.

In oceanic areas, such as western Britain and western Ireland, ombrotrophic nutrition is only slightly inferior to weak minerotrophic nutrition, and little difference in vegetation results. The Boreal zone is largely subcontinental or continental, and ombrotrophic nutrition is much poorer over these areas, particularly with regard to the salts derived from the sea. With decreasing concentrations of mineral nutrients in the precipitation, the nutritional differences between ombrotrophic and minerotrophic peatlands naturally tend to become more marked.

On the other hand, Boreal soils and 'run-off' are such that much of the terrestrial water received by peatland has been in contact only with very poor and leached soils, or has been diluted by large quantities of almost completely ombrotrophic water. As a result, Boreal lands contain much peatland that is only weakly or even doubtfully minerotrophic. This is particularly so within the areas resting on the two large shields of hard Archaean or Pre-Cambrian rocks, the Canadian Shield and Fennoscandia.

Other areas show richer minerotrophic nutrition. Where fragments of limestone are represented in the soils, minerotrophic sites will be strongly enriched in calcium (but often not in other nutrients), and calcareous alkaline, or at least near-neutral, peats are formed. Elsewhere most, though not all, minerotrophic peats are more or less acid, some of them nearly as much so as ombrotrophic peats.

However, acidity is not a simple function of mineral supply. Among other factors influencing acidity, the rate of peat formation seems to be important. In waterlogged sites, a low rate usually results in less acid conditions.

The plant growth-rate and the organic production by photosynthesis are slower in the north, with its short growth-season; in addition the rate of decay of the peat in waterlogged sites seems to be faster in the north. The higher solubility of oxygen in water at low temperatures may be of importance, but most probably the vigorous growth of algae in these waterlogged sites is the main factor. The algae compete strongly with the peat-forming mosses, and they also produce much oxygen by photosynthesis, so that the organic material formed decays easily and no new peat accumulates. Moreover, remains of other plants, as well as existing peat, will be affected by a very active type of wet-surface oxidation, which occurs in addition to the normal decay. It has been shown [11] that there has been no peat formed for 2000–3000 years in some waterlogged areas, called 'hollows' and 'flarks', in the central Swedish uplands.

PRINCIPAL PLANT COMMUNITIES

Much has been written on the composition and ecology of the vegetation of Boreal peatland [9, 10, 12]. There are only a few plant species that are able to grow under strictly ombrotrophic conditions, and these plants also enter most minerotrophic sites; the majority of peatland plants, when growing in the Boreal zone, are limited to minerotrophic sites. Even in weakly minerotrophic (and strongly acid) peat areas, the flora is considerably richer than in ombrotrophic areas, and includes more species of sedges (*Carex*), cotton-grasses (*Eriophorum*), and peat-mosses (*Sphagnum*), and also the buck-bean (*Menyanthes trifoliata*), several willows (*Salix*), and other species. Near-neutral peatland and, still more markedly, alkaline calcareous peatland, have so many plants that these localities have been termed 'rich fen' [10].

The difference in vegetation between ombrotrophic and minerotrophic areas affects the

general colour of the peatland, and they may sometimes be distinguished in aerial colour-photographs (figure 3). On black-and-white photographs their identification is not possible from the tone, but the patterns reveal so much of the hydrography that ombrotrophic and minerotrophic vegetation is usually readily distinguished. Other variations exist in Boreal peatland vegetation [6, 7, 9, 12]. They are effects of, for example, geographical and climatic differences, or of degrees of firmness and wetness.

Typical of Boreal peatland are extremely wet and muddy areas with much algal vegetation but scanty macro-vegetation. They are frequently very extensive but nearly always sharply circumscribed, and are regularly covered by shallow water after the melting of the snow and during other wet periods. They form the most distinctive part of the patterns, and seem to be stable and persistent. This has been shown by pollen stratigraphy [11]. Only rarely do they become filled in with any vigorously peat-forming vegetation. The growth of algae and the 'corrosive" superficial oxidation may explain this lack of recolonization and formation of new peat.

The patterns that occur in Boreal peatlands fall into two groups, those where permafrost does not occur, and those where it does. These groups can be considered separately.

PATTERNS IN AREAS WITHOUT PERMAFROST

In ombrotrophic bogs, ancient hollows that are 'left behind' tend to become deeper in relation to the surrounding bog surface when the peat there grows in thickness; finally they become deep secondary bog-pools [11, 13]. Bog-pools, usually from two or three to about ten feet deep, occur in large numbers in some Boreal bogs, such as in Sweden [11, 13] (figures 7 and 8) and in Quebec and northern Ontario [5] (figure 2). They are not confined to Boreal bogs. Recently, fine examples, evidently with a similar history, have been described from the Silver Flowe in south-west Scotland [14]. Other examples of patterned peatland outside the Boreal zone exist in northern Minnesota [15]. Within the Boreal, by far the most extensive and striking patterning is in the minerotrophic peatland, because of the comparatively large amounts of surface and subsurface run-off.

The principal feature of the Boreal minerotrophic patterns is the 'flark' (in Swedish) or 'rimpi' (in Finnish) (figures 5, 10). The French-Canadian term 'mare' has also been proposed [16]. A typical flark is a limited, usually elongated, wet, and muddy area accessible only with difficulty on foot, because the network of root-stocks often has too coarse a mesh to support the explorer's weight. Each flark is exactly horizontal, even when it is several hundred metres in length. On sloping sites the flarks are narrow, being only a few metres wide, but in almost horizontal peatland they may be a hundred metres wide or more.

The flark vegetation is remarkably uniform all over the Boreal zone, because the principal plants are circumboreal: there will be species of *Carex*, such as *C. limosa*, *C. chordorrhiza*, and *C. livida*, and also *Eriophorum angustifolium*, *E. gracile*, *Menyanthes trifoliata*, *Scheuchzeria palustris*, *Equisetum fluviatile*, *Juncus stygius*, and a few other species. There is still no reasonable explanation why these plants, which all propagate by branching rhizomes, grow so sparsely in the flarks. Nutrition is poor, but seemingly not worse than in other parts of the peatland, where the vegetation grows densely. All these plants are well provided with internal aerating tissues and do not suffer from the lack of oxygen at root level.

The long axis of the flarks is always at right angles to the slope, that is, in the direction of the contours. This is, in fact, true of all patterning that is independent of permafrost or strong winds. On a slope, even if slight, the consecutive flarks form steps, as in a staircase. Some of the large groups of flarks have a striking resemblance to a terraced rice-field (figure 5). In addition, the flark complexes are very clearly limited to the parts of the peatland through which most of the water has to pass on its way downhill.

Both these conditions show that the flarks are primarily dependent on the water flow across the peatland. Several other proposed explanations are ruled out by these basic facts of orientation and situation. It is not likely that downhill movements or soil-creep are the cause, for the flarks would then be found mainly on strong slopes and would be absent from nearly horizontal areas. In fact they are not. On the contrary, the most conspicuous flarks are found on slopes that are much too small to allow for any peat-flow at all (figures 5, 10), and there are no flarks on the steeper slopes (figure 9).

There is evidence [11] that peat growth is slightly better between the flarks than within them. This is necessary for the seemingly precarious stability of the flark system. Another prerequisite is that the damming rims should not be eroded. Indeed, natural erosion is exceptional, being observed only on slopes unusually steep for

flark formation. When it occurs it destroys the pattern completely.

The flarks tend to be very large nearer the sub-Arctic parts of the Boreal zone. Here the dams separating them are frequently, though not always, marked ridges, usually two to three feet high, with lichens, dwarf-shrubs, and occasional low trees growing on them. The reason for the increase in height seems to be the pressure of the ice, exerted horizontally during the freezing period in early winter [17]. Similar phenomena occur in lake ice and are known to pile up stones and soil to form a low ridge round the lake. The expansion of the ice is due to temperature changes; when the ice shrinks in cold weather, numerous narrow cracks are formed which fill with water and then freeze over almost instantly. When the temperature rises again, the ice will expand and exert a pressure on the edges of the basin. The still unfrozen ridges between the flarks will become elevated because of the pressure exerted on them by the expanding ice on the flarks. The ridges become frozen later than the flarks, and they remain frozen much longer in spring.

Even if the pressure of the ice is the cause of the upheaval of the ridges it is probably not the cause of flark formation, nor of patterning as a general phenomenon, as the flarks are also often separated by much lower rims. Although the hydrodynamic pseudo-stability is not a full explanation of the origin of the flark patterns, their relation to hydrodynamic conditions must be taken into account in any hypothesis that is formulated.

PATTERNS IN AREAS WITH PERMAFROST

The first permafrost to appear in the peatlands is located in certain dry hummocks, which through the action of permafrost become increasingly elevated. Several types of elevated and frozen peat mounds are known and are called, *inter alia*, spruce islands and palsas.

Spruce islands are found in Alaska and the interior of Canada (figure 1). At the southern margin of the area of discontinuous permafrost, where the climatic conditions for tree growth are still good, the islands are often quite extensive, though moderately high, and covered by unusually dense stands of black spruce (*Picea mariana*) (figure 4). This species is well adapted to grow on peaty permafrost hillocks. The amount of soil it needs for rooting, and its need of nutrients for growth, are both low; it endures high acidity and reproduces profusely by 'layering', that is by rooting branches that become embedded in litter or moss and lichen vegetation.

Further north, the permafrost hillocks support only scattered low trees (black spruce and *Larix laricina* in Canada). They are blown free of snow in winter and are fully exposed to the cold air. Similar mounds, often entirely treeless, also occur in the northernmost parts of Russia and Fenno-scandia [9, 18], where they are known as palsas (figure 11). They are usually of small size but may be up to 23 ft in height in the extreme north of Sweden and Finland, and still higher in Russia. Between the palsas, freezing is only seasonal, and the ground is usually very wet in summer.

The interior of palsas consists of alternating sheets of frozen peat and pure ice [18], and it is believed that water accumulates on the growing ice-crystals. This water, derived from below or from the surroundings, adds to the bulk of the palsa. A palsa frequently becomes eroded and ultimately may entirely collapse, and is then replaced by a water-filled pit (figure 12).

Where permafrost is continuous, the palsas are replaced by less conspicuous, but very extensive, elevated peat-plateaux, showing frost polygons and other types of pattern, which are very different from the patterns in the Boreal peatlands that are without permafrost. Occasional high mounds of ice and frozen peat occur also in areas of continuous permafrost, provided that liquid water is available. Near the Arctic these 'pingos' are very imposing; the famous pingos in the delta of the Mackenzie River [19] reach the dimensions of the Albert Hall in London.

HISTORICAL EVIDENCE

Only a small fraction of the Boreal peatlands have originated from lakes filled in by vegetation. However, such spots probably often served as starting points for the spread of peatlands over mineral soils. Other early centres were probably the result of the activity of beavers, which were at one time very numerous in these areas. Their dams cause waterlogging on the sides of streams and small rivers, which kills off forest trees and introduces peat-forming vegetation.

Most of the Boreal peats rest directly on mineral soil that once carried forest or at least swampy woodland or scrub. In fact, it seems unlikely that much peatland existed in the Boreal in the early part of the post-Glacial period, except perhaps in non-glaciated areas. Having regard to the recent origin of Boreal peatland and of its plant communities (particularly of the acidophilous

types), their great uniformity in widely separated areas is astonishing. It is likely that most of the species are very old, but their combination into plant communities must be comparatively recent, so far as the present Boreal zone is concerned. The Boreal peatlands spread rapidly in the middle post-Glacial period. Later, this process became slower, and there is now little danger, in Fennoscandia, of deforestation by swamping [20].

Most of the patterns seem to have been lacking during early phases of development of the peatland and have probably originated during more recent millennia. Pollen dating [11] shows that the patterns are about 2000–3000 years old in the central Swedish uplands, and similar ages have been found for sub-Alpine palsas [18] in northernmost Sweden; more precise data may now become available by radiocarbon dating.

Succession in peatland includes much more than the normal successions starting in a wet habitat that are described in ecological textbooks. Obviously the reverse trend is more frequent, in which the development starts from mineral soil and the site gradually becomes wetter, with some parts becoming flarks or even secondary pools. Indeed, a further development seems to occur occasionally where adjacent pools expand and join to form lakes, as in some parts of the Hudson Bay lowland (figure 6). Structures in areas with permafrost also show both building and degradation.

The development of the northern peatlands is hardly in accordance with a uniform 'climax' theory. Rather than leading to uniformity, the paludification and further changes have increasingly diversified the habitats and their vegetation. In particular, the formation of patterns contradicts any climax concept, even a more modest 'polyclimax' view.

A biological material, in itself tending to disperse at random, is forced into a pattern of high mechanical precision by the ecological action of external forces and internal transformation. The variation is endless in detail, yet there is a surprising approximation to what looks like design but, of course, is not. Needless to say, biologists are puzzled, but also fascinated.

REFERENCES

[1] LINNAEUS, C. 'Iter lapponicum'. Several Swedish editions, all posthumous. English edition by Sir James Smith, 1811 ('Lachesis Lapponica'). See also JAMES, W. O., *Endeavour*, **16**, 107, 1957.

[2] HARE, F. K. *Geog. Rev., New York*, **40**, 615, 1950; *Geog. Studies, London*, **1**, 4, 1954; *Canada, Dept. Mines Tech. Surv., Geog. Branch*, mem. **6**, 1959.

[3] LAVRENKO, E. M. and SOCHAVA, V. B. (Ed.). Растительный покров СССР (Vegetational cover of the USSR). 2 vols., being the text of Геоботаническая Карта СССР (Geobotanical Map of the USSR), by the same editors (1954). Academy of Sciences of the USSR, Moscow and Leningrad. 1956.

[4] BROWN, R. J. E. *Arctic*, **13**, 163, 1960.

[5] COOMBS, D. B. *Geog. Bull., Ottawa*, **6**, 1, 1954.
HUSTICH, I. *Acta Geog., Helsingf.*, **16**, Pt. 1, 1957.
SJÖRS, H. *Arctic*, **12**, 2, 1959; *Bull. Nat. Mus. Can.*, **171**, 1, 1961; *ibid.* (In preparation).

[6] DRURY, W. H. *Contr. Gray Herb. Harv.*, **178**, 1956.

[7] ALLINGTON K. R. *McGill Sub-Arctic Res. Pap., Montreal*, 7, 1959 (Thesis). See also [2].

[8] GALKINA, E. A. In [3], Vol. II, p. 553.

[9] RUUHIJÄRVI, R. *Ann. bot. Soc.* '*Vanamo*', *Helsingf.*, **31**, Pt. 1, 1960.

[10] DU RIETZ, G. E. *Svensk bot. Tidskr.*, **43**, 274, 1949; *Vegetatio, Haag*, 5–6, 571, 1954.

[11] LUNDQVIST, G. *Sveriges geol. Unders.*, **Ca21**, 1951.

[12] CAJANDER, A. K. *Acta for. fenn.*, **2**, 1913.
MALMSTRÖM, C. *Medd. Skogsförsöksanst. Stockh.*, **20**, 1923.
KOTILAINEN, M. J. *Wiss. Veröff. Finn. Moorkulturver.*, **7**, 1928; *ibid.*, **19**, 1951.
PAASIO, I. *Acta. for. fenn.*, **39**, 1933.
SJÖRS, H. *Ark. Bot.*, 33A, Pt. 6, 1946; *Acta phytogeogr. Suec.*, **21**, 1948; *Bot. Notiser*, 1950, 173, 1950; *Oikos*, **2**, 241, 1950.
GORHAM, E. *Quart. Rev. Biol.*, **32**, 145, 1957.

[13] GRANLUND, E. *Sveriges geol. Unders.*, **C 373**, 1932.

[14] RATCLIFFE, D. A. and WALKER, D. *J. Ecol.*, **46**, 407, 1958.

[15] H. E. WRIGHT and M. L. HEINSELMAN. Unpublished information.

[16] HAMELIN, L.-E. *Cahiers Géog., Québec*, **2**, 87, 1957.

[17] AUER, V. *Acta for. fenn.*, **12**, Pt. 2, 1920.

[18] LUNDQVIST, G. *Geol. Fören. Stockh. Förh.*, **73**, 209, 1951.

[19] PORSILD, A. E. *Geog. Rev.*, **28**, 46, 1938.

[20] MALMSTRÖM, C. *Medd. Skogsförsöksanst. Stockh.*, **26**, 1, 1931.

Positronium

F. F. HEYMANN

Positronium, a very short-lived complex of an electron and a positron, has some properties resembling those of hydrogen, and can in some respects be looked upon as the lightest of the elements. For example, it reacts with halogens to form halides akin to those of hydrogen, and for each line in its optical spectrum there is a corresponding one at half the wavelength in the spectrum of hydrogen. Positronium closely links physics and chemistry, requiring the techniques of both for its study. This article describes the history and present state of the research.

POSITRONS AND ELECTRONS

The Schrödinger wave equation, which specifies the wave-like behaviour of an elementary particle, is applicable only to cases in which the velocity of the particle is negligible in comparison with that of light. One of the most striking predictions that resulted when P.A.M. Dirac, in 1930, developed a wave equation in which proper account was taken of the consequences of the Special Theory of Relativity, was that of the existence of a positively charged electron. Two years later, a cloud chamber track of this particle, now known as the positron, was indeed observed by C. D. Anderson.

From Dirac's theory, and the theory of Quantum Electrodynamics, it was deduced, and has since been amply verified observationally, that electrons and positrons can be created in pairs in a variety of suitable circumstances, and, conversely, that an encounter between a positron and an electron may lead to the annihilation of both particles, the energy associated with their masses and motion being converted into electromagnetic radiation in the form of γ-rays.

The electron and the positron are both spinning particles. Each may be thought of as an electric charge rotating about its own axis in a top-like fashion. A natural consequence of this rotation is that both particles possess a magnetic field, that of a positron being directed oppositely to that of an electron spinning in the same sense. The inherent angular momentum of these particles is referred to as their spin.

Spin, being an angular momentum, is subject to certain well-established quantum rules. One important consequence of these rules is that, when an electron and a positron are close to one another, their spin axes can be only either parallel or anti-parallel to one another. These two possible states of affairs are known as the triplet and singlet states respectively, and will be referred to frequently in this article.

By invoking the principle of conservation of angular momentum, it can be shown that when a positron and an electron annihilate one another in a head-on collision in the singlet state, all the initial energy is converted into two γ-rays emitted in directions 180° apart when measured in the C.M. co-ordinates (co-ordinates in which the centre of mass of the initial particles is at rest). If the particles move relatively slowly before the collision, which is so in all the cases to be discussed here, the energy involved is substantially just that associated with the rest-masses of the two particles. In numerical terms, each γ-ray will then have an energy of 0·511 MeV.

In the case of a head-on annihilating collision in the triplet state, it can be shown that three γ-rays must be produced. These three γ-rays are coplanar in the C.M. co-ordinates and the sum of their energies is 1·022 MeV, but the energy of any one of them is statistically distributed in a continuous spectrum as illustrated in figure 1. On the same diagram the singlet-annihilation radiation would be a line spectrum at 0·511 MeV.

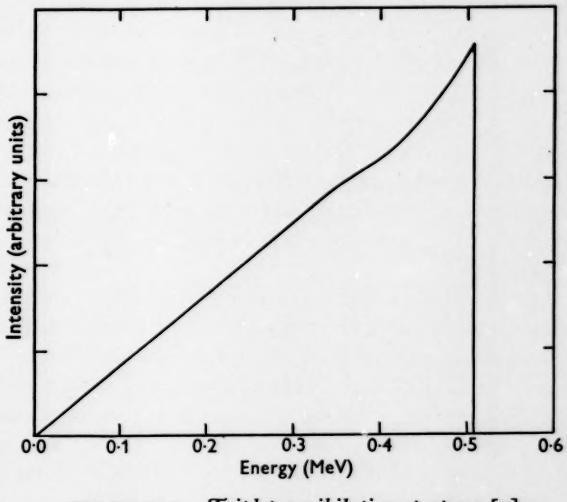


FIGURE 1 – *Triplet-annihilation spectrum* [5].

Glancing collisions hardly ever result in annihilation; the two cases described above are the only ones of practical importance.

The probability of a triplet-annihilating collision is only about 1/1120 of that of a singlet annihilation. This means that the number of triplet annihilations makes an almost negligible contribution to the rate at which positrons are annihilated in free collisions when moving through a cloud of electrons with randomly oriented spin-axes. Nevertheless, the smallness of the triplet-annihilation probability is the very feature which makes a study of positronium at all possible, as will be shown later.

THE NATURE OF POSITRONIUM

In the preceding discussion of the annihilation of a positron and an electron it has been tacitly assumed that the two particles meet in a free head-on collision leading to their disappearance. Although the ultimate fate of every positron is an act of annihilation with an electron, it is indeed fortunate for the enquiring mind that the path to death of a positron frequently occurs through a more devious route.

It is, for instance, possible for a positron moving in a gas to prise loose one of the electrons of a gas molecule and to remain attached to it. In classical language, this electron-positron pair rotates around their common centre of mass, held in dynamic equilibrium by a balance between their electrostatic attraction and the centrifugal force due to the rotary motion. It is tempting to describe such a structure as an atom; and there is a close analogy between it and the hydrogen atom in which an electron and a proton, nearly two thousand times heavier than the positron, rotate about their common centre of mass. The electron-positron pseudo-atom is what is known as positronium.

It should be stressed that this formation of bound states is by no means a rare phenomenon. It plays a major part in the study of positron annihilation. Experimental evidence is available to show that, typically, 30 per cent of all slow positrons in a gas form positronium atoms before annihilating.

Depending on the original relative orientations of the electron and positron spin-axes, positronium can be formed in either the triplet or the singlet state. From statistical arguments it can be shown that formation in the triplet state is three times as likely as in the singlet state.

Although positronium is a stable structure in the sense that it requires energy to separate the particles from one another, the positron and the electron nevertheless eventually suffer mutual annihilation, and have a predictable mean life. In the case of the singlet state, this is in fact about $1{\cdot}25 \times 10^{-10}$ seconds, with the usual inverse exponential distribution of individual lifetimes. As was mentioned previously, annihilation in the triplet state is less likely by a factor 1/1120, leading in this case to a mean life of about $1{\cdot}4 \times 10^{-7}$ seconds. These times may sound negligibly short, but they are quite long on an atomic scale. To be sure, the mean life of singlet positronium is too short to allow observations to be made on more than a few of its properties with present-day techniques; but 10^{-7} seconds is long enough to study many aspects of the physical and chemical behaviour of triplet positronium.

Positronium has a system of excited states, the energies of which can be calculated in exactly the same way as that of the hydrogen atom. In fact, apart from differences in the fine structure of the levels, primarily due to the differences in mass and magnetic moment of protons and positrons, the energy levels of positronium are similar to those of atomic hydrogen, with all energies scaled down by a factor which is almost exactly 2. The ionization energy—the energy required to just separate the positron and the electron from unexcited positronium—is consequently 6·8 eV, or half that of the hydrogen atom. Moreover, for each line in the optical spectrum of hydrogen there is a corresponding line in the positronium spectrum at double the wavelength.

QUENCHING OF TRIPLET POSITRONIUM

Once formed, the positronium atom will make random collisions with the gas molecules in which it moves, frequently leading to chemical and quasi-chemical processes. Evidence has, for instance, been cited for the existence of such chemical compounds as halides of positronium.

The effect usually observed as arising from interactions between positronium and gas molecules is an apparent reduction in the mean life of triplet positronium by processes collectively known as quenching collisions.

Probably the most frequently occurring quenching mechanism is one in which, during a collision between a positronium atom and a gas molecule, the positron annihilates directly with one of the electrons of the molecule in a singlet orientation. At first sight, one is led to suggest that this would yield a rate of annihilation comparable with free positrons moving in the same gas. In the case of

positronium, however, the positron is screened to a large extent by the electron, resulting in a reduction of the annihilation rate by about one order of magnitude. In gases in which this is the major process, one would therefore expect to find a moderately small degree of quenching.

Two groups of gases are known to be markedly strong quenching agents. The first group, notably NO and NO_2, contains an odd number of electrons in the molecule. The process responsible for quenching in this case is thought to be one in which, during a collision, the electron of the triplet positronium is exchanged for the odd electron in the molecule, which may be in a suitable orientation to result in singlet positronium. The singlet state then annihilates rapidly with its own characteristic mean life of $1{\cdot}25 \times 10^{-10}$ seconds.

The second group of strongly quenching gases, namely the halogens, are thought to form chemical compounds such as positronium chloride, analogous to hydrochloric acid. In such a compound, instead of wandering freely from molecule to molecule, the positronium is bound to remain close to the chlorine electrons, thus greatly enhancing the probability of an annihilating collision.

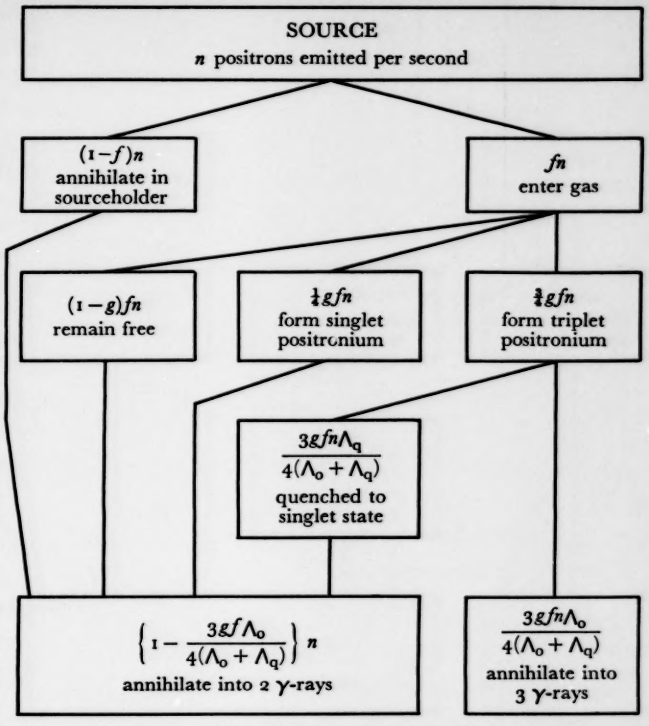


FIGURE 2 – *Possible fates of positrons in a gas* [6].

FATE OF POSITRONS IN A GAS

The positrons utilized for studies on positronium are usually obtained by placing a radioactive source such as ^{22}Na, which decays by emitting an energetic positron, in the gas under observation. The positrons are rapidly slowed down by collisions with gas molecules until they move slowly enough either to annihilate or to form positronium. This slowing-down time is comparatively short (about 10^{-9} seconds) in comparison with both the mean life of triplet positronium and the mean life of positrons against annihilation in free collisions with electrons.

A chart summarizing the possible alternative courses of the history of a positron in a gas is shown in figure 2. In this diagram the factor f represents the fraction of the emitted positrons that manages to leave the source and enter the gas without annihilating with electrons of the material of the source or source-holder. The factor g represents the fraction of the positrons entering the gas that forms positronium.

The quantity Λ_o is the natural annihilation rate of triplet positronium, that is, the reciprocal of the mean life in the absence of quenching; while Λ_q is the quenching rate, that is, the reciprocal of the mean life if quenching were the only factor causing disappearance of triplet positronium.

DIRECT OBSERVATION OF POSITRONIUM LIFE IN A GAS

It is one of those rare fortunes of nature that the radioactive material commonly used for studying positronium emits a 1·36 MeV γ-ray practically simultaneously with the emission of the positron. The detection of this γ-ray thus provides a convenient time marker for registering the instant of birth of a positron. The instant of annihilation can be observed accurately by detecting any one of the annihilation γ-rays. A sketch of typical equipment used for observations of this kind is shown in figure 3. R is the radioactive source, placed inside a high-pressure vessel C. A scintillating crystal X_1 detects the passage of a 1·36 MeV γ-ray, thus defining the time of birth of a positron. This crystal is optically coupled to a photomultiplier P_1 which converts the light flash into an electrical impulse and amplifies it. A lead plug L prevents radiation from the source from reaching the crystal X_2, which can, however, respond to radiation produced by a positron annihilating in the gas. This crystal, in conjunction with the

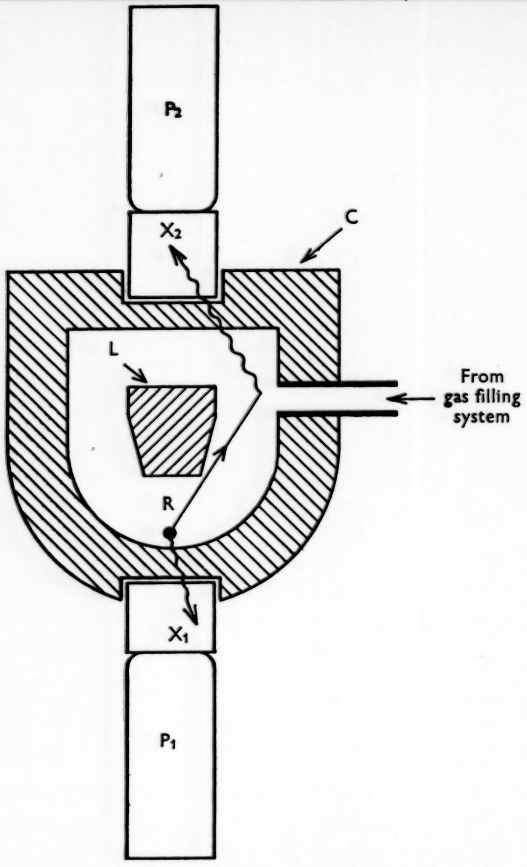


FIGURE 3 – *Apparatus for observing positron life.*

photomultiplier P_2, produces a pulse to define the instant of annihilation. The time intervals between individual 'birth' and 'death' pulses are measured and recorded electronically. With suitably fast electronic circuits it is possible to measure these time intervals to an accuracy of about 10^{-9} seconds, which is around 1 per cent of the mean life of triplet positronium.

In experiments of this kind, positrons that form singlet positronium immediately after slowing down annihilate so rapidly that, because of the finite time-resolution of the equipment, their lifetimes show up as effectively prompt coincidences between the birth and death counters. Positrons annihilating by free collisions with electrons and triplet positronium do, however, usually yield measurably long lives. This is illustrated by the typical delayed coincidence curve of figure 4. This curve can, for instance, be analysed into three components as shown, of which the slow exponential component, manifesting itself as a straight line on a logarithmic plot, contains all the information necessary to determine the mean life and hence its reciprocal, the annihilation rate.

Lifetimes arising from free annihilations may be distinguished experimentally from those due to triplet positronium by virtue of the fact that their annihilation rates differ in their dependence on pressure. In the case of free annihilation, positrons will collide with electrons of the gas molecules at a rate proportional to the mean electron-density. In other words, the annihilation rate under these circumstances will be proportional to the gas pressure. Results obtained from observations in oxygen (figure 5) illustrate this state of affairs. Oxygen is such a strong quenching agent that, even at small pressures, triplet positronium has such a short life that the only significant contribution to the long-lived component is that due to annihilation of free positrons. The curve of annihilation rate against pressure is accordingly a straight line passing through the origin, corresponding to a strict proportionality.

In contrast to this, gases of the Freon series, typified by CCl_2F_2, have the remarkable property of mopping up free positrons and causing them to annihilate rapidly, leaving triplet positronium as the only contributor to the long-lived component. In the absence of quenching, the mean life of triplet positronium is a quantity characteristic of the positronium atom alone, independent of gas pressure. This is indeed found to be the case in Freon, as shown in figure 5.

In practice, the long-lived component usually has to be ascribed to a mixture of free positron, and triplet-positronium, annihilations. The life curve under these circumstances will consist of a mixture of two exponentials. The two parts may be

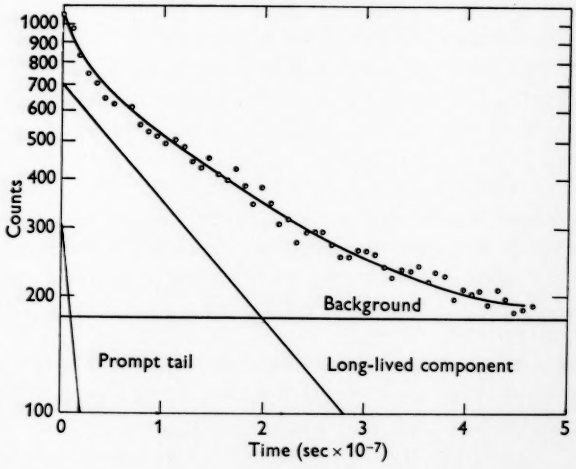


FIGURE 4 – *Delayed coincidence curve in Freon* [4].

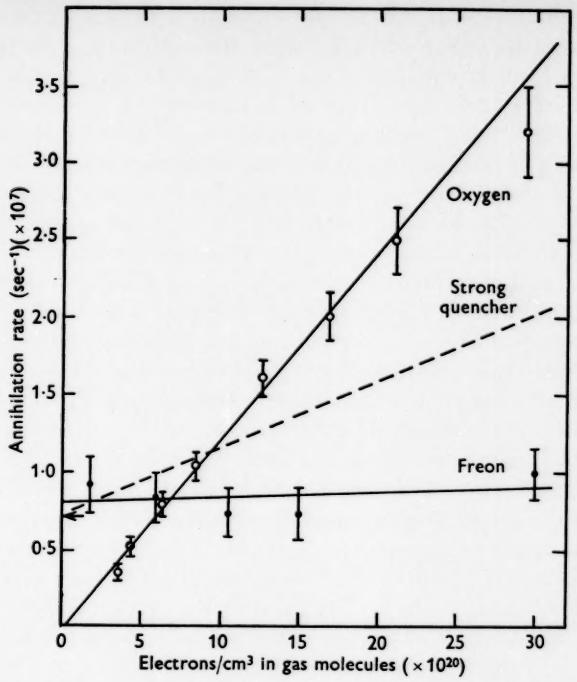


FIGURE 5 – *Pressure-dependence of annihilation rates in oxygen and Freon* [3]. *Arrow indicates theoretical triplet-rate* [5].

observed separately by adding either a small amount of Freon to the gas under observation in order to remove the free positrons, or a small amount of a strong quenching agent such as NO in order to remove the triplet positronium.

The third line in figure 5 illustrates the effect of a gas with a strong quenching action. The variation of lifetime with pressure is such that the annihilation rate consists of two additive components; a constant rate due to the natural mean life, with a quenching rate proportional to electron density superimposed on it.

γ-RAY-SPECTRUM OBSERVATIONS

As was shown earlier (figure 1), the annihilation γ-rays for triplet annihilations have a continuous energy-spectrum, while the spectrum of singlet annihilation consists of a sharp line. This difference between the two spectra is used to determine the effective concentration of triplet positronium in a gas sample.

When a beam of mono-energetic γ-rays, such as those resulting from singlet annihilations, impinges on a scintillating crystal such as sodium iodide, the resulting electrical pulses from a photomultiplier coupled to the crystal have a height distribution as shown in figure 6, curve (i). The relatively sharp peak is produced by γ-rays

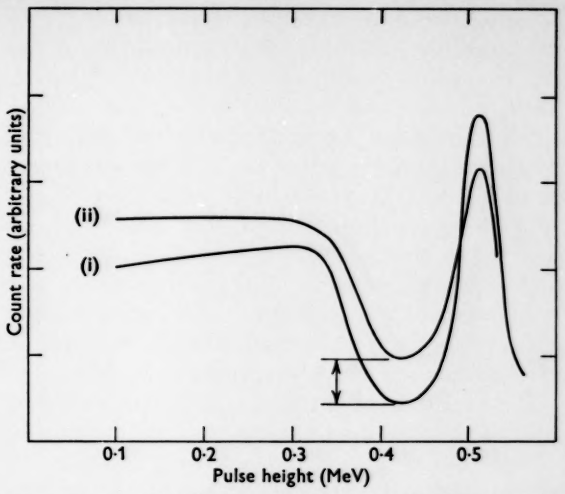


FIGURE 6 – *Scintillation pulse-height spectra from annihilation radiation.*

suffering photoelectric absorption in the crystal, in which case the magnitude of the light pulse is directly proportional to the energy of the γ-ray. The long plateau to the left arises largely from those γ-rays that are scattered by Compton collisions with electrons, in which case a smaller and variable fraction of the γ-ray energy is converted into light. Because of this complex response of a scintillation spectrometer to even a mono-energetic beam of γ-rays, it is in general not practicable to retrieve detailed information about the energy distribution in a continuous range of γ-ray energies from the resulting pulse-height distribution. Nevertheless, if such a unit is allowed to view the γ-rays produced in a chamber where both singlet and triplet annihilations are occurring, the resulting pulse-height spectrum is modified to that shown in figure 6, curve (ii). The increase in the valley height, shown by the arrows, is a sensitive indication of the number of three-quantum annihilations.

What is perhaps the simplest demonstration of the existence of triplet positronium may therefore be made by comparing the valley height arising from annihilation radiation in a chamber containing a positron source and a weakly quenching gas with that which results when a small admixture of NO is introduced into the chamber. In the latter case the triplet state is strongly quenched, so that practically all the annihilation γ-rays are in the form of two-quantum radiation.

OTHER DETECTION METHODS

The remaining important methods that have been used for identifying the origin of annihilation

radiation make use of the facts that the multiplicity and the angular distributions of triplet radiation differ from those of singlet radiation.

For instance, two scintillation-counters viewing a chamber from opposite directions will give coincident output pulses as a result of singlet annihilations. This has been used indirectly to observe triplet positronium, by observing the decrease in the intensity of two-quantum radiation when conditions in the chamber are altered to result in more three-quantum annihilation, such as by doing successive observations with and without small admixtures of NO to the gas in the chamber.

The emission of three simultaneous γ-rays in triplet annihilations has been verified directly by placing three scintillation-counters with their centres coplanar with the centre of the positronium chamber, in positions such that no straight line through any two crystals intersects the chamber. Threefold coincidence-pulses between these counters then indicate the occurrence of a triplet annihilation. By this means the coplanarity of the radiation has been verified by observing the reduction in counting rate when one of the counters is displaced from the coplanar position.

Some of the more sophisticated observations on positronium have used more than one of the detection techniques simultaneously. For instance, spectral measurements on pure triplet positronium have been made in which the pulse-height distribution of one of the scintillation counters in a triple counter arrangement was analysed, counting only those pulses accompanied by coincident pulses from the other two.

MAGNETIC QUENCHING

When a triplet positronium atom is placed in a magnetic field, the interaction between the field and the magnetic moments of the electron and the positron may cause the spin axis of one or the other particle to be reversed in direction for part of the time. In the absence of other effects there is then a distinct probability, depending on the strength of the magnetic field, of finding singlet positronium in what was initially a sample of triplet positronium. In the language of quantum mechanics, the presence of a magnetic field causes mixing to occur between the amplitudes of the state with unit angular momentum and a magnetic quantum number of zero (a certain fraction of the triplet positronium) and of that with zero angular momentum (singlet positronium). During that fraction of the time in which the original triplet atom behaves like a singlet atom there is, of course, a greatly enhanced probability that it will annihilate rapidly, thus effectively leading to a reduced mean life for the triplet state. An equivalent way of stating this is to say that, in a magnetic field, triplet positronium can annihilate in the two-quantum mode characteristic of the singlet state, the magnetic field itself absorbing any unbalance in angular momentum necessary to satisfy the conservation laws.

If a positron source is placed in a gas sample, the concentration of triplet positronium in it will reach a dynamic equilibrium when the rate of formation from newly emitted positrons is balanced by the rate of loss due to annihilation. A reduction in the mean life will then lead to a reduction in the concentration. This fact has been used to observe magnetic quenching, using the valley-height method of measuring the intensity of the triplet-annihilation spectrum. A typical curve plotting the rate of triplet annihilations against magnetic field is shown in figure 7. The observations are entirely in agreement with detailed theoretical calculations.

ENERGY LEVELS AND OPTICAL SPECTRUM

If triplet positronium were formed in an excited state, there would be a distinct probability of emission of one of its characteristic optical spectral lines if the mean life against de-excitation by optical emission were short in comparison with the mean life against direct annihilation in the excited state. In particular, this is the case for that line analogous to the Lyman α line of atomic hydrogen. In the positronium spectrum, this line has a wavelength of about 2430 Å. Although this is twice the wavelength of the hydrogen line, it is

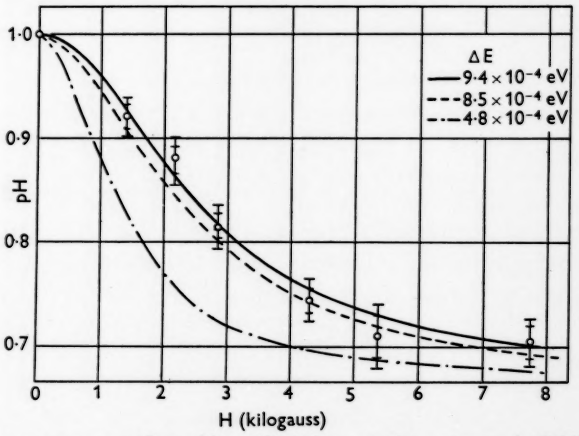


FIGURE 7 – *Quenching of triplet state by a magnetic field* [2]. *The various curves correspond to different theoretical interpretations.*

still well in the ultra-violet region, which makes for some complication in its detection. In addition, the concentrations of positronium obtainable in practice are extremely small in comparison with gas-concentrations typically used for spectroscopic work, so that one would not expect to detect more than one photon every few seconds in a typical laboratory experiment. Although a number of attempts have been made to locate this line, no completely unambiguous result has yet been obtained.

In spite of the fact that no optical line in the positronium spectrum has yet been observed, one is confronted with the experimental situation that certain details of the fine structure of the lines due to the Zeeman effect have been observed.

In a magnetic field, spectral lines in general are found to be split into components differing very slightly in wavelength. This is because the magnetic field causes various energy-levels of the atom to be split into close-lying components, corresponding to different values of the magnetic quantum number. By irradiating the atoms with a radio-frequency electromagnetic field superimposed on the static magnetic field, it is possible to induce transitions between these sub-levels when the frequency is that corresponding to the differences in energy involved.

In a noteworthy set of experiments of M. Deutsch and S. C. Brown, this process was in fact observed in the following manner. A positronium chamber in the shape of a microwave-resonant cavity was placed in a static magnetic field strong enough to ensure practically complete magnetic quenching. In these circumstances, as was mentioned earlier, almost all the triplet positronium with magnetic quantum number zero would be quenched. A radio-frequency field was then applied to the cavity, whereupon further quenching was observed, at that value of the static field which caused the correct level-splitting to enable the radio-frequency field to induce transitions from the sub-levels with magnetic quantum number ± 1 to the sub-level with magnetic quantum number zero. This additional quenching, plotted against magnetic field, is shown in figure 8.

The position of the minimum in this curve agrees to a high accuracy with the value calculated from quantum electrodynamics when a number of refined theoretical arguments are included.

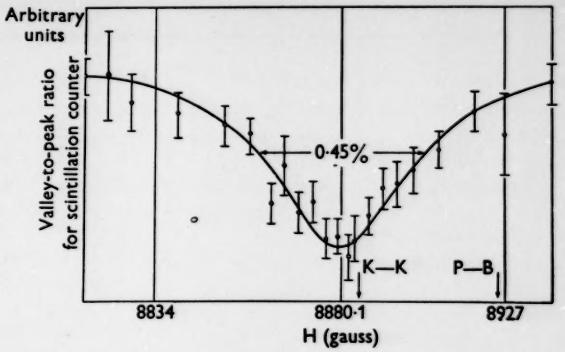


FIGURE 8 – *Quenching of $m = \pm 1$ level by radio-frequency field* [1]. *P—B is expected minimum position according to first-order electrodynamic theory. K—K is expected minimum position including higher-order corrections.*

This was a remarkable experiment by any standards, and is to be compared with the famous Lamb-Retherford experiment on the ground-state-level splitting of the hydrogen atom, as being one of the few precision experiments that have been performed to prove the correctness of certain abstruse higher-order effects in electrodynamics.

CONCLUSION

In an article of this length it has not been possible to do more than touch on some of the more salient features of the study of positronium. The author does, however, hope to have indicated that the detailed study of the fate of a positron in matter has proved to be a subject with many more facets of interest than was originally expected from the simple Dirac theory of annihilation by head-on collisions between positrons and electrons, in so far as this quasi-atomic complex of a transitory nature known as positronium may be formed.

The fascination of the experimental study of the behaviour of positronium lies perhaps in the particular way that it allows the nuclear physicist to contribute to chemistry and spectroscopy, using his now traditional techniques of scintillation counting. In this field microchemistry has in one sense been carried to its ultimate limit, in that the reaction times of individually reacting atoms have been observed.

REFERENCES

[1] DEUTSCH, M. and BROWN, S. C. *Phys. Rev.*, **85**, 1047, 1952.
[2] DEUTSCH, M. and DULIT, E. *Ibid.*, **84**, 601, 1951.
[3] DEUTSCH, M., DULIT, E., and PRIMAKOFF. *M.I.T.L.N.S.E. Prog. Rep.*, 131, 31st May, 1951.
[4] DUFF, B. G. Private communication.
[5] ORE, A. and POWELL, J. L. *Phys. Rev.*, **75**, 1696, 1949.
[6] OSMON, P. E. Thesis, University of London, 1960.

Book reviews

THEORETICAL PHYSICS

Theoretical Physics in the Twentieth Century: a memorial volume to Wolfgang Pauli, *edited by M. Fierz and V. F. Weisskopf. Pp. x+328. Interscience Publishers Inc., New York; Interscience Publishers Ltd, London.* 1960. *72s. net.*

This collection of twelve articles, to which has been added a bibliography of all his known published work, was originally intended as a sixtieth-birthday gift to Pauli: instead, as a result of his death before it was complete, it became a memorial volume. What emerges from these accounts of twentieth-century physics is that there is scarcely any aspect of it, except perhaps molecular structure, where Pauli's work has not been decisive. It was in the study of fundamental problems that his main interest lay: here, imagination and courage are needed, as well as technical ability. Pauli possessed them all. And the results of this are the Exclusion Principle, the neutrino, the Pauli spin matrices, the general formulation of quantum electrodynamics, and much else. Some of the historical chapters, particularly those by Kronig and Van der Waerden, provide admirable accounts of topics of central importance, and they show, far better than a series of more formal papers, the curious way in which fundamental discovery really proceeds. Most of the book is understandable by any scientist with some training in the physical sciences. The majority of the chapters are in English, but three are in German.

C. A. COULSON

THERMODYNAMICS

Monographs in Statistical Physics and Thermodynamics, *Vol. II,* Thermodynamics, *by P. T. Landsberg. Pp. x+499. Interscience Publishers Inc., New York; Interscience Publishers Ltd, London.* 1961. *£5 9s. net.*

The author of this book is Professor of Applied Mathematics at University College, Cardiff, which helps to explain his approach to thermodynamics as 'among the most abstract branches of Physics'. While this book undoubtedly fulfils the author's aim, which is 'to supplement existing texts, and contribute something to our understanding of thermodynamics', it would not be suitable for a newcomer. Thermodynamics is too useful a science to be omitted, even in elementary courses, but it requires very careful introduction. In this book the treatment is clearly recognized by the author as unconventional. In the first of no fewer than seven appendices, Professor Landsberg attempts to define some of the more unusual features in his book, including the use of the concepts of 'diathermic' and 'adiabatic' partitions. Many interesting, and quite a few important, questions are tucked away in the six other appendices and in the problems presented at the end of each chapter. For advanced reading, the book contains much that gives access in an original way to familiar aspects of thermodynamics, though the degree of originality and unconventionality of treatment are not in all cases justified by the economy of method or profundity of results achieved. A course of reading in this book, for those already well grounded in statistical thermodynamics developed in a more orderly way, is likely to be valuable for theoretical physicists, and would enlarge their experience.

A. R. UBBELOHDE

FLAME PHOTOMETRY

Flame Photometry, *by J. A. Dean. Pp. viii+341. McGraw-Hill Book Co. Inc., New York; McGraw-Hill Publishing Co. Ltd, London.* 1960. *89s. net.*

This book is one of a series on advanced chemistry, and is written primarily for those using flame photometry as an analytical method, for whom it can be strongly recommended.

The data on the various elements, which will be of principal value to the intended reader, are comprehensive both in number and detail. Much of the material is new, and the presentation is systematic and orderly. Of particular interest is Professor Dean's own work on organic-solvent extraction of metals before estimation by flame photometry, and on the increased sensitivity obtained with organic solvents in the flame. The gain in sensitivity is thought to be due partly to the resultant increase in flame size. This points to inadequate filling of the monochromator, but the appropriate remedy is not suggested.

Certain criticisms can be made, mainly in the chapters devoted to instrumentation. Effects attributed to the magnitude of the slit width are somewhat vague. A table (9-1) intended to illustrate the percentage increase in flame background with increased slit width appears to show the the converse. Other anomalies tend to mar these sections of the book but are unlikely to perplex the general reader, who should find much of value and interest.

R. L. WARREN

TRANSPORT PHENOMENA

Transport Phenomena, *by R. B. Bird, W. E. Stewart, and E. N. Lightfoot. Pp. xxi+780. John Wiley & Sons Inc., New York and London.* 1960. *£5 10s. net.*

This book is based on notes issued to students in the chemical engineering course at the University of Wisconsin. It deals with the transport of momentum, energy, and mass—that is, matter—in fluids. The analogies between these processes, both in the formal and physical senses, are pressed as far as they will go, although there is of course no analogy in the case of radiative-energy transport, which is also included.

The theme is logically developed and rigorously treated, and in these respects is a long way ahead of most of the texts for chemical engineers in this field. It is quite uncompromising in its approach, and far from easily readable. It would, I think, need the backing of an interpretative set of lectures if used as a text for undergraduates.

The current tendency for the engineering syllabus in U.S. universities to 'go scientific' is here reflected in an avoidance of topics, such as heat transfer to boiling liquids, that cannot be dealt with analytically. I believe this trend may go too far, and that we ought to produce more powerful teaching methods for developing both insight and the qualitative analysis of problems. Meanwhile, this is a first-rate book of its kind.

P. V. DANCKWERTS

AMMONIA SYNTHESIS

La Sintesi dell'Ammoniaca, *by C. A. Vancini. Pp. xxvii+1075. Editore Ulrico Hoepli, Milan.* 1961. *Lire 12 000.*

Total world production of nitrogen compounds, mainly derived from synthetic ammonia, now exceeds twelve million tons of equivalent nitrogen per year, and is still expanding rapidly. Many variations in synthesis-plant design have been introduced since ammonia was first synthesized commercially, and a wide variety of processes are now in use for the production

and purification of the vast amounts of hydrogen required. This book sets out the technical and scientific background of these processes in a comprehensive fashion. New processes for hydrogen production, such as steam re-forming of methane and partial oxidation at pressure, are included, as well as the long-established processes, such as those based on coke, coke-oven gas, and electrolysis of water. The recently developed purification processes, such as CO_2 removal with hot potassium carbonate solutions, are included, except, surprisingly, the Italian 'Vetrocoke' process. Low-temperature processes for the production of oxygen, nitrogen, and mixtures of nitrogen and hydrogen are also given. Design of pressure vessels and of ammonia-synthesis plant generally is well covered. Each chapter has a comprehensive bibliography of technical and scientific literature, with, however, few references to patents. The book is well provided with line diagrams, and the theoretical parts are illustrated with simple calculations. Managers and designers of ammonia plants, or those concerned with large-scale hydrogen production, will find it a most useful book. An English translation would be very welcome. T. J. P. PEARCE

ORGANIC ANALYSIS

Organic Electronic Spectral Data. *Vol. I* (1946–52), *edited by M. J. Kamlet, pp. xiv*+1208. *Vol. II* (1953–55), *edited by H. E. Ungnade, pp. x*+919. *Vol. I, £10 15s. net; vol. II, £6 12s. net. Interscience Publishers Inc., New York; Interscience Publishers Ltd, London.* 1960.

The visible and ultra-violet spectra of organic substances have been studied very extensively in recent years. The theoretical interpretation of the spectra, and especially the empirical correlation of the spectra with molecular structure, have been of great value to the organic chemist. The volume of published spectra has grown enormously, as convenient and accurate photoelectric spectrophotometers have become available; a group of over fifty chemists has started to prepare a compilation of all available spectra.

Volumes I and II cover the years 1946–52 and 1953–55, and the editors and their collaborators are to be congratulated on undertaking their enormous task. These two volumes contain a mass of data that will be of great value to all organic chemists.

The compounds are arranged according to molecular formulae, as in 'Chemical Abstracts'. The tables give the compound name, the solvent used, the wavelengths of the absorption maxima, the extinction coefficient of each band, and a reference to the original paper. In Volume I alone the tables cover 1160 pages.

R. E. RICHARDS

QUINONES

Ciba Foundation Symposium on Quinones in Electron Transport, *edited by G. E. W. Wolstenholme and C. M. O'Connor. Pp. xii*+453. *J. & A. Churchill Ltd, London.* 1961. 60*s. net.*

The discovery, by Professor R. A. Morton and his collaborators at Liverpool University, and independently by Professor D. E. Green's group at the University of Wisconsin, of compounds of the ubiquinone (coenzyme Q) series, has provided fresh stimulus to the study of the roles of lipids in electron transport and in the associated phosphorylation processes in mitochondria. We are indebted to the initiators and organizers of this symposium, which took place in 1960, for providing the means whereby representatives of many groups of investigators could report and discuss the initial discoveries, the elucidation of chemical structures, and the many ramifications which are now being pursued.

Apart from authoritative contributions directly related to the title, aspects of the subject which received attention include *in vitro* chemical mechanisms of oxidative phosphorylation; possible roles of quinones (including vitamins K) and of related chroman derivatives in phosphorylation; the significance of plastoquinone (Kofler's quinone) in photosynthetic electron transport and phosphorylation; the biosynthesis of the polyisoprenoid and quinoid portions of the ubiquinones; and incidental critical appraisal of techniques employed in these investigations.

The scope of the symposium demonstrates the attraction of the field for a variety of specialists, and this book will be valuable to an equally wide circle of readers. M. B. THORN

BIOCHEMISTRY

Metabolic Pathways, *Vol. I, edited by D. M. Greenberg* (*second edition of* Chemical Pathways of Metabolism). *Pp. xv*+572. *Academic Press Inc., New York; Academic Press Inc.* (*London*) *Ltd, London.* 1960. £6 1*s.* 6*d. net.*

This book should be of great interest to those—and there must be many—who wish to consult a fully documented reference book that surveys existing knowledge on metabolic pathways in animals, plants, and micro-organisms. This second edition has been virtually rewritten. Of the fourteen contributors to Volume I, only four were among the authors of the first edition. In the seven years since the first publication, knowledge of intermediary metabolism has grown extremely rapidly—many gaps have been filled and many new and unexpected metabolic sequences have been discovered. The subjects covered include free energy and entropy in metabolism (A. B. Pardee and L. L. Ingraham); mitochondrial system of enzymes (D. E. Green and S. Fleischer); glycolysis (B. Axelrod); the tricarboxylic acid cycle (H. A. Krebs and J. M. Lowenstein); other pathways of carbohydrate metabolism (B. Axelrod); metabolism of steroid hormones (L. T. Samuels); and carotenoids and vitamin A (G. Mackinney).

H. A. KREBS

INSECT BIOCHEMISTRY

The Biochemistry of Insects, *by D. Gilmour. Pp. xii*+343. *Academic Press Inc., New York; Academic Books Ltd, London.* 1961. 64*s. net.*

This is the first book to have been written on insect biochemistry and its appearance is timely, because this field has recently become very much more active. The author set out to describe the chemistry of insects as organisms, rather than the chemistry of reactions, and he has achieved this by relating the various biochemical processes to the peculiarities of insect physiology and biology. The result is a very readable book.

Well over half the book is devoted to various aspects of energy production, and in these chapters we see clearly how most of the processes are similar to those of other animals, but with important differences of detail, as in glycolysis and the respiratory chain. The remainder of the book is principally concerned with the intermediate metabolism of carbohydrates, lipids, and nitrogen compounds, including surveys of the components of insect cuticle and of the theories of cuticle tanning.

A number of important papers have been overlooked (for example, the nutrition of mosquitoes is described from early papers, not from the later definitive work), and the references are listed in a manner which could hardly

be less convenient for the reader; but the book as a whole provides a most useful critical survey of insect biochemistry. A. N. CLEMENTS

BOTANY

Botanisk Atlas, *Vol. II, by O. Hagerup and V. Peterson (English translation by H. Gilbert Carter). Pp.* 299. *Ejnar Munksgaard, Copenhagen.* 1960. *D.kr.*76.

This is a companion volume to volume I, by the same authors, which was reviewed in ENDEAVOUR in July 1957. It extends the work to the conifers and vascular cryptogams, and is carried out on the same general plan and with the same merits as its precursor. The omission to state magnifications of the drawings that was noted in the previous volume has been made good in this, to the great increase of their usefulness. About half the Danish species of mosses and liverworts are illustrated, and all those of the lycopods, horsetails, and conifers. As before, there are habit drawings (the trees are represented only by small shoots), usually from fresh material, and enlarged drawings of detail where required. There is rather more accompanying text than in the first volume. It consists mainly of a description of each of the groups illustrated, and is given in Danish and English in parallel columns. Finally there is a section of phylogenetic speculation, with particular reference to the origin of the seed habit, which seems a little at variance with the factual nature of the main body of the work. It is to be regretted that in the illustrations of some groups, such as the Sphagnales, the usual clarity of outline has given place to a fuzziness which leaves a good deal to the imagination. Nevertheless, with the completion of this second volume the authors have increased the debt that botanists owe them. W. O. JAMES

INVERTEBRATE ZOOLOGY

Traité de zoologie, *edited by P.-P. Grassé. Vol. V, Part* 2: Bryozaires, Brachiopodes, Chétognathes, Pogonophores, Mollusques. *Pp.* 1168. *Masson et Cie, Paris.* 1960. *Bound, NFcs* 190; *paper covers, NFcs* 180.

This volume, the second of three concerned with the invertebrates between Platyhelminthes and Nematodes on the one side and Arthropoda on the other, maintains the very high standard of technical excellence set by this magnificent treatise, which now comes near to completion. Both text-figures and plates are models of clarity.

The Bryozoa are very fully covered in a chapter of almost 300 pages by Paul Brien, followed by an account of their fossil representatives by Emil Buge. This is followed by chapters on recent and fossil Brachiopoda, by Paul de Beauchamp and Jean Roger respectively, in which recent work on feeding currents and embryology is well summarized, and the structure, classification, and evolutionary history of the group are thoroughly described. The section on the Chaetognatha, also by Paul de Beauchamp, appears somewhat brief, and certainly much more could have been made of the great ecological importance of these animals in the plankton of all seas.

The chapter of some 100 pages on the relatively recently established group of the Pogonophora is of particular value. It is by A. V. Ivanov of the University of Leningrad, the recognized authority on these animals, and one of the few contributors who is not French. This is the best general account available, and has the basic structure displayed, semi-diagrammatically, in two admirable coloured plates.

The remainder of the volume, somewhat over 500 pages, deals with the Molluscs, excluding Gastropoda, Scaphopoda, and Cephalopoda, which will form the subject matter of the third volume. A. Portmann deals with the general characters of the phylum, and exception might be taken to his failure to note the different symmetries of the body and of the mantle that are the key to an understanding of the wide variety of form within the Mollusca. E. Fischer-Piette and A. Franc deal very adequately with Aplacophora and Polyplacophora, while the Monoplacophora are covered by H. Lemche and K. G. Wingstraad, joint authors of the recent description of *Neopilina*. The value of some of their functional inferences must await examination, unlikely to be long delayed, of living specimens. P. Manigault deals briefly with the shell, and then A. Franc covers the great class of the Bivalvia. Writing as one whose work, with that of co-workers, looms large in this chapter, I can say that modern work has been well summarized and well illustrated. There is certainly no better account available. While the continued use of Pelseneer's classification can hardly be defended, the large systematic section contains excellent descriptions of superfamilies and families. Finally, fossil bivalves are well covered, by Colette Dechaseaux.

All interested in any of these groups should obtain, or make sure they have access to, this fine book.

C. M. YONGE

CERAMIC SCIENCE

Oxide Ceramics, Physical Chemistry and Technology, *by E. Ryshkewitch. Pp. viii*+472. *Academic Press Inc., New York; Academic Press Inc. (London) Ltd, London.* 1960. $16.

The information this book contains could be presented much more clearly in about one-third of the number of pages used in the present volume. For example, a paragraph of nine lines is found necessary to define Young's modulus, and this paragraph includes the statement 'Perhaps the most significant strength magnitude of a solid body is its Young's modulus of elasticity'. Similar inexactitudes may be found on nearly every page.

The author is not writing in his native language; this book is based on his *Oxydkeramik der Einstoffsysteme*, published in 1948. It is a great pity that someone could not have been found to help improve the style of writing, for there appears to be much useful information to be extracted from this volume. The first sixty-eight pages are a general introduction to oxide ceramics; then follow chapters on alumina (150 pages), spinel (23 pages), magnesia, beryllia, zirconia, and thoria and ceria, and a short section on oxide-metal and oxide-carbide ceramics. The author index is comprehensive and the subject index is adequate. The book appears to be addressed primarily to those concerned with the development of articles made from oxide-ceramics, and as a source of data and a guide to the literature it will be useful. There are many useful hints on the preparation of oxide-ceramic articles; it is in this practical field that the author has made noteworthy contributions.

R. W. DOUGLAS

Short notices of books

(These notices are descriptive rather than critical and are designed to give a general indication of the nature and scope of the books.)

Quantum Mechanics, *by J. L. Powell and B. Crasemann. Pp.* x+495. *Addison-Wesley Publishing Co. Inc., Reading, Mass.* 1961. 74*s. net.*

This is an introduction to quantum mechanics intended for senior students at universities. It aims at emphasizing the physical basis of the subject while not ignoring the mathematics. There are, *inter alia*, chapters on the origins of the quantum theory, the Schrödinger equation, operators and eigenfunctions, matrix mechanics, perturbation theory, and identical particles.

Progress in Low Temperature Physics, *Vol. III, edited by C. J. Gorter. Pp. xii+* 495. *North-Holland Publishing Co., Amsterdam.* 1961.

Volume IV, like the first two, is a collection of review articles intended for research workers. Some of the chapters are: vortex lines in liquid helium II, by W. F. Vinen; the nature of the λ-transition in liquid helium, by M. J. Buckingham and W. M. Fairbank; recent developments in superconductivity, by K. W. Taconis; solid-state masers, by N. Bloembergen; and some solid-gas equilibria at low temperatures, by Z. Dokoupil.

A Reference Book of Chemistry, *by J. H. White. Pp.* 302, *University of London Press Ltd, London.* 1960. 21*s. net.*

This book is intended as a supplement to the textbooks used in the higher forms of schools. The sections include: definitions and explanations of important terms, named organic reagents and tests, named inorganic reagents and tests, and named organic reactions. Appendices deal with first aid in the laboratory and with 'useful data'.

Principles of Physical Chemistry, *by W. H. Hamill and R. R. Williams, Jr. Pp. x+*607. *Oliver and Boyd Ltd, Edinburgh.* 1960. 30*s. net.*

The treatment used in this book is to develop thermodynamics first, and then chemical physics; it is intended for use in universities. The whole of fundamental physical chemistry is covered, with chapters on, among other subjects, the first law of thermodynamics, chemical equilibrium in gaseous systems, entropy and the third law of thermodynamics, phase diagrams, nuclear chemistry, and photochemistry and radiation chemistry.

An Introduction to Co-ordination Chemistry, *by D. P. Graddon. Pp. vii+* 111. *Pergamon Press Ltd, Oxford.* 1961. 25*s. net.*

The advances of the last two decades are covered by this book, which also includes a brief chapter on the early history of co-ordination chemistry. There are also chapters on polymerization and co-ordinate saturation, the stability of complex salts, and carbonyls and π-complexes. The book is intended primarily for students.

Handbuch der Papierchromatographie, *Vol. II,* Bibliographie und Anwendungen, *edited by I. M. Hais and K. Macek. Pp. xxiv+*726. *Gustav Fischer Verlag, Jena.* 1960. *DM* 44.

This second volume of an extensive treatise on paper chromatography consists of a comprehensive classified bibliography of 10 290 items, referring mainly to the period 1943–56. The book contains, firstly, a general part that covers history, general reviews and books, the theory and techniques of chromatography and preparative chromatography, and, secondly, a special part containing sections on the chromatography of groups of compounds, for example the phenols and related substances, organic acids, steroids. The sections themselves are further subdivided into techniques and applications and are preceded by brief introductory paragraphs of text. The copious indexes fill nearly 200 pages.

Modern Chemical Processes, *Vol. VI, by the editors of 'Industrial and Engineering Chemistry' in conjunction with the technical staffs of the co-operating organizations. Pp.* 126. *Reinhold Publishing Corporation, New York; Chapman and Hall Ltd, London.* 1961. 48*s. net.*

This series aims at describing the processes of American chemical technology in terms of their principles, their plant, their economics, and their future. Some of the techniques described in this volume are: the manufacture of basic silicone products, the use of Grignard reagents, the manufacture of specialized surface-active agents, and the manufacture of acrylates and methacrylates.

A Bibliography of Eastern Asiatic Botany, *by E. H. Walker. Pp. xl+*552. *American Institute of Biological Sciences, Washington, D.C.* 1960. $18.50 *to individuals and industrial libraries,* $16.50 *to AIBS members and all other libraries.*

The original Bibliography, compiled by E. D. Merrill and E. H. Walker, was published in 1938, and covered the literature to the end of 1936. The present volume extends the coverage to the end of 1958 and is arranged in the same way as the original volume. The author explains that Japan and, more markedly, China have not been completely covered, because of the wartime destruction of some libraries in Tokyo and because of the poor circulation of publications in China during the war.

Non-convertible Coatings, *by The Oil and Colour Chemists' Association. Pp.* 326. *Chapman and Hall Ltd, London.* 1961. 30*s. net.*

The primary purpose of this book is to help students to prepare for the Non-convertible Coatings Section of the City and Guilds Examination in Paint Technology, although it is hoped that it will be useful to those in the industry. It is the first of six parts. The contents include chapters on nitrocellulose, plasticizers, lacquers, cellulose acetate, synthetic thermoplastic lacquer resins, and spirit varnishes.

The Major Achievements of Science, *by A. E. E. McKenzie. Vol. I, pp. xvi +*368. *Vol. II, pp. xii+*195. *Cambridge University Press, Cambridge.* 1960. *Vol. I,* 30*s. net* (*School edition* 20*s. net*); *Vol. II,* 17*s.* 6*d. net* (*School edition* 12*s.* 6*d. net*).

The author of this book has aimed at giving students of the humanities a background of science. He gives a survey of scientific discoveries, in chapters such as 'The circulation of the blood', 'Field physics', and 'The germ theory of disease', and of the historical background, in chapters devoted to the seventeenth, eighteenth, nineteenth, and twentieth centuries. The patterns of scientific thought, and the characters of the scientists, are emphasized. The second volume consists of extracts, translated where necessary, from the writings of distinguished scientists.

Notes on contributors

A. F. WAGNER
A.B., M.A., Ph.D.
Was born in Jersey City, New Jersey, in 1922 and was educated at Princeton University, from which he received a Ph.D. degree in 1951. He joined the research staff of Merck & Co. Inc. after leaving the university. His main interests are in the synthesis, and the relationship between structure and activity, of natural products; he participated in the first effective synthesis of mevalonic acid.

K. FOLKERS
B.S., Ph.D.
Was born in Decatur, Illinois, in 1906 and was educated at the University of Illinois, and then at the University of Wisconsin, where he was awarded a Ph.D. in 1931. From 1931 to 1934 he was a Post-Doctorate Research Fellow at Yale University under Professor Treat B. Johnson. He joined Merck & Co. Inc. in 1934; since 1956 he has been Executive Director of Fundamental Research. His interests have been in the chemistry of therapeutic agents, such as alkaloids, vitamin B6, and pantothenic acid; his more recent studies have resulted in the discovery of mevalonic acid, and he has lately been investigating the chemistry of the coenzyme Q group. He has received the Perkin Medal of the Society of Chemical Industry, and is President-elect of the American Chemical Society.

SIR EDWARD BULLARD
M.A., Sc.D., F.R.S.
Was born in 1907 and was educated at Repton School and Clare College, Cambridge. He began research under Lord Rutherford in 1929, working on the scattering of slow electrons in gases. In 1931 he took up geophysics and has done experimental work on gravity, seismology, and the measurement of heat flow on land and at sea. From 1950 to 1955 he was Director of the National Physical Laboratory, and is now Reader in Geophysics in the University of Cambridge. He has written extensively on the origin of the Earth's magnetic field and on magneto-hydrodynamics.

M. H. PIRENNE
M.A., Ph.D., Dr.Sc.
Was born in Verviers, Belgium, in 1912, and was educated at the University of Liège. From 1942 to 1945 he was seconded to Sir F. C. Bartlett's laboratory in Cambridge for research on vision. From 1945 to 1948 he was an I.C.I. Research Fellow at the University of London, from 1948 to 1955 Senior Lecturer in Physiology at the University of Aberdeen. Since 1955 he has been at the University of Oxford, where he is now a University Demonstrator in Physiology. In 1957 he was elected a Vice-President of the *Comité International de Photobiologie*. His main interests in research are the physiology of vision, and in subjects related to this.

V. A. EYLES
D.Sc., F.R.S.E., F.G.S.
Was born in Bristol in 1895 and was educated at Fairfield Secondary School and Bristol University. After service with the Royal Engineers and Royal Air Force (1915–19), he was appointed geologist to the Geological Survey of Great Britain (Scottish Branch) in 1920 and transferred to the English branch as District Geologist in 1945, a post held until his retirement in 1955. He is a founder member of the British Society for the History of Science and member of the History of Science Society (U.S.A.). He has contributed papers on the history of geology to a number of periodicals, and lectured on the subject in the Department of the History and Philosophy of Science, University College, London, from 1951 to 1956, and also at the Universities of California and Illinois.

H. SJÖRS
Fil. dr.
Was born in Sweden in 1915 and was educated at Uppsala University, where he became Lecturer in Plant Ecology in 1948. He later spent three years at the University of Lund, and since 1955 has been a *laborator* in forest botany at the School of Forestry, Stockholm. His main work has been on the ecology of bog and fen vegetation. He has also published works on the phytogeography of Scandinavia, on humus soils, and on northern woodlands conditioned by ancient types of rural management, and is actively interested in nature conservation.

F. F. HEYMANN
B.Sc.(Eng.), Ph.D.(Eng.)
Was born in South Africa in 1924 and was educated at Capetown University from 1940 to 1945. He was on the staff of the University until 1947, when he came to England and joined Metropolitan-Vickers as an industrial research physicist. Since 1950 he has been on the staff of University College, London, where he is a Reader in Physics. He has published extensively on particle accelerators and high-energy nuclear physics; for the last three years he has been engaged in research on positronium.